

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
ptnr:SP TREMBL- ACC:Q9D398	6330415E02RIK PROTEIN - Mus musculus (Mouse)	945	862/945 (91%)	897/945 (94%)	0.0
ptnr:SP TREMBL- ACC:O08722	TRANSMEMBRANE RECEPTOR UNC5H2	945	862/945 (91%)	893/945 (94%)	0.0
ptnr:SP TREMBL- ACC:O08747	UNC-5 HOMOLOG (C. ELEGANS)	931	610/929 (65%)	723/929 (77%)	0.0
ptnr:SP TREMBL- ACC:O95185	TRANSMEMBRANE RECEPTOR UNC5C - Homo sapiens	931	598/929 (64%)	718/929 (77%)	0.0

10

15

- 20

25

30

35

40

45

5	NOV1b	ENSKWSACSTECAHWRSRECMAPPPGCGRDCSGTLDDSNCTDGLCNC	-----TBAGSEALYAGLVVAVF	374
	Q9D398	ENSKWSACSTECAHWRSRECMAPPPGCGRDCSGTLDDSNCTDGLVLRQRTLNDPKSHPLETSCVVALYAGLVVAVF	-----TBAGSEALYAGLVVAVF	386
	008722	ENSKWSACSTECAHWRSRECMAPPPGCGRDCSGTLDDSNCTDGLVLRQRTLNDPKSRPLEFSGVALYAGLVVAVF	-----TBAGSEALYAGLVVAVF	386
	008747	SNKWSACSTECATWRRRECFATAPNCGKDCGLVLCSSNCTDGLCNC	-----APDSSEVALYVGVVAVF	389
	095185	ENSKWSACSTECATWRRRECFATAPNCGKDCGLVLCSSNCTDGLCNC	-----APDSSEVALYVGVVAVF	389
10	NOV1a	VVALLMAVGCVVYRRNCRDFDITDSSAALTCGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYRCGPVYALQCSADKI	-----	454
	NOV1b	VVALLMAVGCVVYRRNCRDFDITDSSAALTCGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYRCGPVYALQCSADKI	-----	454
	Q9D398	VVALLMAVGCVVYRRNCRDFDITDSSAALTCGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYRCGPVYALQCSADKI	-----	466
	008722	VVALLMAVGCVVYRRNCRDFDITDSSAALTCGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYRCGPVYALQCSADKI	-----	466
	008747	CLATLMAVGCVVYRRNCRDFDITDSSAALTCGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYRCGPVYALQCSADKI	-----	463
15	NOV1a	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	534
	NOV1b	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	534
	Q9D398	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
	008722	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
	008747	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
20	NOV1a	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	533
	NOV1b	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	533
	Q9D398	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
	008722	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
	008747	PMTNPSLLDPLPSLRILKVYSSSTGSGCGLAGDADLLGVLPCTYPCDPSRTHFLHLRSASLGSQLLGLPRDPSSSVS	-----	546
25	NOV1a	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	614
	NOV1b	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	614
	Q9D398	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
	008722	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
	008747	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
30	NOV1a	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	613
	NOV1b	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	613
	Q9D398	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
	008722	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
	008747	GTFCGLGGRLLPTCTGCVSLVLLGCAIPGKGFYEMYLINNAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTHPHCA	-----	626
35	NOV1a	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	694
	NOV1b	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	694
	Q9D398	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
	008722	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
	008747	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
40	NOV1a	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	693
	NOV1b	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	693
	Q9D398	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
	008722	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
	008747	EVSARDWIFOLKTOAHQGHWEVVTLEETLNTPCYCOLPBRACHILLDQLGTYVFTGESYSRSRAVKRLQIAFPALCT	-----	706
45	NOV1a	SLEYSRLVYCLEDTFVALKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	774
	NOV1b	SLEYSRLVYCLEDTFVALKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	774
	Q9D398	SLEYSRLVYCLEDTFVALKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
	008722	SLEYSRLVYCLEDTFVALKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
	008747	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
50	NOV1a	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	773
	NOV1b	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	773
	Q9D398	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
	008722	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
	008747	SLEYSRLVYCHIDTADLKEVLELERTLGGYVLEEPHPLFKDSYHNRLSLHCHIDPAHWRSKLLAKYOEIPFYHWSGS	-----	786
55	NOV1a	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	854
	NOV1b	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	854
	Q9D398	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	866
	008722	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	866
	008747	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	852
	NOV1a	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	852
	NOV1b	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	852
	Q9D398	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	866
	008722	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	866
	008747	QNALHCTFTFLERHSLASTELTCKVCRVQVEGEGIFQLHITTLAETPAGSLDACSAPGSTVITOLGPYAFKIPLSIROKI	-----	852
	NOV1a	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	933
	NOV1b	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	933
	Q9D398	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	945
	008722	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	945
	008747	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	931
	NOV1a	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	931
	NOV1b	CSSLDAPOGRCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEAQDQDGDLSLASALEENGKSEMLVATDQGD	-----	931

The presence of identifiable domains in NOV1, as well as all other NOVX proteins, was determined by searches using software algorithms such as PROSITE, DOMAIN, Blocks, Pfam, ProDomain, and Prints, and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (<http://www.ebi.ac.uk/interpro>). DOMAIN results for NOV1 as disclosed in Tables 1G-1O, were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST analyses. This BLAST analysis software samples domains found in the Smart and Pfam collections. For Tables 1G-1O and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading or by the sign (!) and "strong" semi-conserved residues are indicated by grey shading or by the sign (+). The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

	NOV2b	-----	1
	NOV2c	KFAFTNLTPFTMYDVYIAAETSAGTGPKSNISVFTPPDVGAVFDLQLAESTOVVRITW	419
5		430 440 450 460 470 480	
	NOV2a	KKPRQPNGIINQYRVKVLVPETGIILENTLLTGNNNEINDPMAPEIVNIVCPMVGLYEGSA	474
	NOV2b	-----	1
	NOV2c	KKPRQPNGIINQYRVKVLVPETGIILENTLLTGNNNEINDPMAPEIVNIVCPMVGLYEGSA	479
10		490 500 510 520 530 540	
	NOV2a	EMSSDLHSLATFIYNSHDPKNFPARNRAEDOTSPVVTRNQYITDIAAEOLTYVHRLRR	534
	NOV2b	-----	1
	NOV2c	EMSSDLHSLATFIYNSHDPKNFPARNRAEDOTSPVVTRNQYITDIAAEOLSYVHRLRP	539
15		550 560 570 580 590 600	
	NOV2a	FWAETMGFSRYTIMSSASRONITSPG---ELSAQNPRVTIVTIDIEVFLWDDPP---PVF	588
	NOV2b	-----	1
20	NOV2c	FTEHMSVSAFTIMGEGPPTVLSVRTRQQVPSIRIINYKNTSSSLLYWDDPPSYPNGK	599
		610 620 630 640 650 660	
25	NOV2a	FHHYLTITLDVENQSKSIILRTLNSLSYVILGLKKYTKYKMRVAASTHVGESSLSEENDI	648
	NOV2b	-----	1
	NOV2c	ITHYLTIVAMELDTNRAFOITITDINSFLITGLGLKKYTKYKMRVAASTHVGESSLSEENDI	659
		670 680 690 700 710 720	
30	NOV2a	FVRTSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGIIAYEVLYKNIDTLYMKNTST	708
	NOV2b	-----	1
	NOV2c	FVRTSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGIIAYEVLYKNIDTLYMKNTST	719
35		730 740 750 760 770 780	
	NOV2a	TDIILRLNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSEIVPDSAPENITYKNISSGEI	768
	NOV2b	-----	1
	NOV2c	TDIILRLNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSEIVPDSAPENITYKNISSGEI	779
40		790 800 810 820 830 840	
	NOV2a	ELSFLPPSSPNGIIQKYTIYLKRSNGNEERTINTTSLTONIKGLKKYTOYIIIEVSASTLK	828
	NOV2b	-----	1
45	NOV2c	ELSFLPPSSPNGIIQKYTIYLKRSNGNEERTINTTSLTONI--LKKYTOYIIIEVSASTLK	837
		850 860 870 880 890 900	
50	NOV2a	GEGVRSAPISILTEEDAPDSPQDFSVKQLSGVTVKLSWQPPLEPNGIILYYTVYVWR--	886
	NOV2b	-----	1
	NOV2c	GEGVRSAPISILTEEDAPDSPQDFSVKQLSGVTVKLSWQPPLEPNGIILYYTVYVWRNR	897
		910 920 930 940 950 960	
55	NOV2a	SSLKTINVTETSLELSDLDYNVEYSAYVTASTRFGDGKTRSNIISFOTPEGSPDPPKDVY	946
	NOV2b	-----	1
	NOV2c	SSLKTINVTETSLELSDLDYNVEYSAYVTASTRFGDGKTRSNIISFOTPEGSPDPPKDVY	957
		970 980 990 1000 1010 1020	
60	NOV2a	YANLSSSSIIILFWTPPSKPNGIIQYYSVYYRNTSGTFMQNFTLHEVTNDFDNMTVSTIID	1006
	NOV2b	-----	1
	NOV2c	YANLSSSSIIILFWTPPSKPNGIIQYYSVYYRNTSGTFMQNFTLHEVTNDFDNMTVSTIID	1017
		1030 1040 1050 1060 1070 1080	
65	NOV2a	KLTIIFYTYFWLTASTSVGNGNKSSDIIIEVYTDQDVPEGFVGNLTYESISSTAINVSWVP	1066
	NOV2b	-----	26
	NOV2c	KLTIIFYTYFWLTASTSVGNGNKSSDIIIEVYTDQDVPEGFVGNLTYESISSTAINVSWVP	1077
70		1090 1100 1110 1120 1130 1140	

5 NOV2a A Q H D G N V T K W Y D A Y F N K A R P Y F T N E G F P N P P C T E G K T K F S G N E E I Y I I G A D N A C M I P G N E 1846
NOV2b A Q H D G N V T K W Y D A Y F N K A R P Y F T N E G F P N P P C T E G K T K F S G N E E I Y I I G A D N A C M I P G N E 806
NOV2c A Q H D G N V T K W Y D A Y F N K A R P Y F T N E G F P N P P C T E G K T K F S G N E E I Y I I G A D N A C M I P G N E 1857

10 NOV2a D K I C N G P L K P K K Q Y L F K F R A T N I M G Q F T D S D Y S D P V K T L G E G L S E R T V E I I L S V T L C I L S 1906
NOV2b D K I C N G P L K P K K Q Y L F K F R A T N I M G Q F T D S D Y S D P V K T L G E G L S E R T L E ----- 855
NOV2c D K I C N G P L K P K K Q Y L F K F R A T N I M G Q F T D S D Y S D P V K T L G E G L S E R T L E I I L S V T L C I L S 1917

15 NOV2a I I L L G T A I F A F A R I R Q K Q K E G G T Y S P Q D A E I I D T K L K L D Q L I T V A D L E L K D E R L T R ----- 1962
NOV2b ----- 855
NOV2c I I L L G T A I F A F A R I R Q K Q K E G G T Y S P Q D A E I I D T K L K L D Q L I T V A D L E L K D E R L T R L L S Y 1977

20 NOV2a P I S K K S F L O H V E E L C T N N N L K F Q E E F S E L P K F L O D L S S T D A D L P W N R A K N R F P N I 2017
NOV2b ----- 855
NOV2c R K S I K P I S K K S F L O H V E E L C T N N N L K F Q E E F S E L P K F L O D L S S T D A D L P W N R A K N R F P N I 2037

25 NOV2a K P Y N N N R V K L I A D A S V P G S D Y I N A S Y I S G Y L C P N E F I A T Q G P L P G T V G D F W R M V W E T R A K 2077
NOV2b ----- 855
NOV2c K P Y N N N R V K L I A D A S V P G S D Y I N A S Y I S G Y L C P N E F I A T Q G P L P G T V G D F W R M V W E T R A K 2097

30 NOV2a T L V N L T O C F E K G R I R C H O Y W P E D N K P V T V F G D I V I T K L M E D V O I D W T I R D L K I E R H G D C M 2137
NOV2b ----- 855
NOV2c T L V N L T O C F E K G R I R C H O Y W P E D N K P V T V F G D I V I T K L M E D V O I D W T I R D L K I E R H G D C M 2157

35 NOV2a T V R Q C N F T A W P E H G V P E N S A P L I H F V K L V R A S R A H D T T P M I V H C S A G V G R T G V F I A L D H L 2197
NOV2b ----- 855
NOV2c T V R Q C N F T A W P E H G V P E N S A P L I H F V K L V R A S R A H D T T P M I V H C S A G V G R T G V F I A L D H L 2217

40 NOV2a T O H I N D H D F V D I Y G L V A E L R S E R M C M V O N L A O Y I F L H O C I L D L L S N K G S N O P I C F V N Y S A 2257
NOV2b ----- 855
NOV2c T O H I N D H D F V D I Y G L V A E L R S E R M C M V O N L A O Y I F L H O C I L D L L S N K G S N O P I C F V N Y S A 2277

45 NOV2a L Q K M D S L D A M E G G D V E L E W E E T T M 2281 (SEQ ID NO:6)
NOV2b ----- 855 (SEQ ID NO:8)
NOV2c L Q K M D S L D A M E G G D V E L E W E E T T M 2300 (SEQ ID NO:10)

50

The disclosed NOV2a polypeptide has homology to the amino acid sequences shown in the BLASTP data listed in Table 2H.

NOV2A

gi|12621078| MDFFHFSFLPLLIGTSESQVDVSSSFDGIGYDILLSSVSA-TTSSSPVSH

gi|125977| -----MGLQMTAARPTAALLLVLSLLLTWTHPTIVDAAPPEIIRK

gi|10728878| -----MGLQMTAARPTAALLLVLSLLLTWTHPTIVDAAPPEIIRK

gi|7290546| -MDCATRKQQQLRAHHKQQQIQIQTHGRKKRQOKQKRHHHHHYVONSOOC

gi|1362625| -MDCATRKQQQLRAHHKQQQIQIQTHGRKKRQOKQKRHHHHHYVONPOOC

60 70 80 90 100

NOV2A

gi|12621078| IGAS--NEPGPPVFLAGERVGSAGILLSMWTPPNPNCGRIISYHYKYKEVC

gi|125977| TLATNVTKPPPPVFLAGERVGSAGILLSMWTPPNPNCGRIISYVVKYKEVC

gi|10728878| PONGVRVGGVASFYCAARGPPPPSVWRKNGKKVSG

gi|7290546| PONGVRVGGVASFYCAARGPPPPSVWRKNGKKVSG

gi|1362625| QKHFWLVVGLLHFLAHANAADLVV--NVPNASSNANAFYRIDYSPPF

QKHFWLVVGLLHFLAHANAADLVV--NVPNASSNANAFYRIDYSPPF

110 120 130 140 150

NOV2A

gi|12621078| PMOTVYTVQVRKPDLSLEVLITNLINPCTIYEIKUAAENAGIGVFSDDPFL

gi|125977| PMOTAYTRARAKPDLSLEVLITNLINPCTIYEIKUAAENAGIGVFSDDPFL

gi|10728878| -T-OSRYTVLEQPGGISILRIEPPVACRADDAPYECVAENGVCDAVSADAT

gi|7290546| GPPPNNTTTPASDIE-KDINFSRALPGTEVNFWYYTNSGTHREQLTWTVN

gi|1362625| GPPPNNTTTPASDIE-KDINFSRALPGTEVNFWYYTNSGTHREQLTWTVN

160 170 180 190 200

NOV2A

gi|12621078| FCTAESAPGKVVVFTGZAVPFSSK-LMWYTS-ATKKYITSEFKISVKHNRS

gi|125977| FCTAESAPGKVVNLIVFALNYSAVNLIIWYLRQPNGKITSEFKISVKHARS

gi|10728878| LTIYE-----GDKTPAGFPVITCGPTRVIEVGHTVLMTCKAIG

gi|7290546| LTIYEGWQ---KTAISGDKTIPAGFPMITCGPTRVIEVGHTVLMTCKAIG

gi|1362625| ITTAPDPF---ANLSVQLRSSKSAFTWRPP--GSGRYGGRIRVRLGLTD

ITTAPDPF---ANLSVQLRSSKSAFTWRPP--GSGRYGGRIRVRLGLTD

210 220 230 240 250

NOV2A

gi|12621078| GIVVREVSIRVFCISASLELHCNENSESEFLWSTASPSPTLGRVTPPSET

gi|125977| GIVVQDVSLRVEDISGKLPEECNENSESEFLWSTTSPSPPTLGRVTPPTVET

gi|10728878| NPTPNYIWIKNQTKVDMSNRYSKDCGFLQENGREEDOCKYECVAENSM

gi|7290546| LPFERYSVLEGNETIOLSAK--ELTFGGSYQOVQYYSVYOGKESVAYTSN

gi|1362625| LPFERYSVLEGNETIOLSAK--ELTFGGSYQOVQYYSVYOGKESVAYTSN

260 270 280 290 300

NOV2A

gi|12621078| THSSSTLTQNEISSV-KEPISFVVTHLRPYTTYLFEEVAVTTEACYIDST

gi|125977| TQSSSTAARSKISSVWKEPISFVVTHLRPYTTYLFEEVAVTTEACYIDST

gi|10728878| CTEHSKATNLIVKVRVPPTFSRPPETISEVMGCSNLNLSCIAVGSPMPH

gi|7290546| CTEHSKATNLIVKVRVPPTFSRPPETISEVMGCSNLNLSCIAVGSPMPH

gi|1362625| FTIKPNTPGKFTVWRPNETILLVWQPPFPAGVITHYRVSITPPDDAIOSV

FTIKPNTPGKFTVWRPNETILLVWQPPFPAGVITHYRVSITPPDDAIOSV

310 320 330 340 350

NOV2A

gi|12621078| NVRTPEVSVPCEGPPONCVTENITGKSSSLLWDPPITVITGKFSYRVELYGPS

gi|125977| NVRTPEVSVPCEGPPONCIMGNTGKASSISWDPPITVITGKFSYRVELYGP

gi|10728878| VKWMKGSSEDLTPENEMPICRNVLQOLINIQESAN--

gi|7290546| NVVERCEGPPCPAAAFKGLMPGNEYNISVQT

gi|1362625| NVVERCEGPPCPAAAFKGLMPGNEYNISVQT

360 370 380 390 400

NOV2A

gi|12621078| AGRILDNSTKDLKFAFTNLTPFTMYDVVYAASTAGTGPKSNYSVFTTPPD

gi|125977| SGRILDNSTKDLRFATHTLTPFTMYDVVYAASTAGVGPKSNYSVFTTPPD

gi|10728878| -----YICIAASTLCQIDSVSVKVQS

gi|7290546| -----YICIAASTLCQIDSVSVKVQS

gi|1362625| -----VSEDEPS-SVPTDARYLTVPER

-----VSEDEPS-SVPTDARYLTVPER

		410	420	430	440	450
	NOV2A	VFCVFDILOLAEVESTOVRTITWKKPRPNGIINQYRVVLVPESTGIITEN				
5	gi 12621078	VFCVFDILOLAEVETELRTITWKKPRPNGIISQYRVVSVLESTGVVTEN				
	gi 125977	LFTAPTDOISEVTATSVRIEWSYKG-----PEDLYYVVICYKPKANAN				
	gi 10728878	LFTAPTDOISEVTATSVRIEWSYKG-----PEDLYYVVICYKPKANAN				
	gi 7290546	VLNVTDEAYITSSG--FRVRNEEPPR--TYSEFDAYQVVLSTSRRTFN				
	gi 1362625	VLNVTDEAYITSSG--FRVRNEEPPR--TYSEFDAYQVVLSTSRRTFN				
10		460	470	480	490	500
	NOV2A	TLITGNNE-INDPMAPEIUNIVQEMVGLYEGSAEMSSDLHSLATFIYNH				
	gi 12621078	TLITGQDEISINPMSPPEIMNLVDEMGIFVEGSGEMSSDLHSPASFIYNH				
	gi 125977	QAFSEISG--IITMYVVRALSEYTEYEFYVIAVNNIGRG-----				
15		510	520	530	540	550
	NOV2A	PDKNFPAARNRAEDQTSPPVITRNQYITDIAAEQLTYVLIRLRRFWAETMG				
	gi 12621078	PHNDFPASTRAEEOSSPPVITRNQYMTDITAEQLSYVVRRLVPFTEHTIS				
	gi 125977	-----PPSAPATCTIGETKMESAP-----				
	gi 10728878	-----PPSAPATCTIGETEMESAP-----				
20		560	570	580	590	600
	NOV2A	FSRYTIMSSASRDNLTPG---PLSACNFRVTHVTIMEVPHMDPPDPV				
	gi 12621078	VSAPTIMGEGPPTVLTVRTREQVPSSIQINYNMISSSSLLYMDPPEYP				
	gi 125977	-----RMQVRTLSSSTVITWEPPETP-----				
	gi 10728878	-----RMQVRTLSSSTVITWEPPETP-----				
	gi 7290546	-----GFLDDG---SNALHISWEPAETG-----				
25		610	620	630	640	650
	NOV2A	FFHHYLIITLDVENOSKS---IILRTLNSLSLVHIGLKKYTKYKRVAA				
	gi 12621078	NGKITHYDIYATELDINR---AFQMTTVDNSFLITGLKKYTRYKRVAA				
	gi 125977	NGQVTGYKVYTTINSNQPEASWNSQMVNDNSLTTSVSDVTPHATYTVRVOA				
	gi 10728878	NGQVTGYKVYTTINSNQPEASWNSQMVNDNSLTTSVSLTPHATYTVRVOA				
	gi 7290546	RQDSYRISYHEQTNASEV---PAPFPVAAESQITTNLTETLDSILAGR				
30		660	670	680	690	700
	NOV2A	STHYGESSLSEENDIFVRTSEDEPESSPDVEVIDVTADEIRLKWSPPEK				
	gi 12621078	STHYGESSLSEENDIFVRTPEDEPESSPDVQVTGVSPSELRLKWSPPEK				
	gi 125977	YTSMCAGPMSTP-----				
	gi 10728878	YTSMCAGPMSTP-----				
	gi 7290546	RYLHAVORLSKG-----				
	gi 1362625	RYLHAVORLSKG-----				
35		710	720	730	740	750
	NOV2A	PNGIIIAEVLYKNIDTLYMKNTSTTDIILRNLRPHPTLYNISVRSYTFPG				
	gi 12621078	PNGIIIAEVLYQNADTLFVKNTSTTDIISDHPKPYTLNISIRSITRLG				
	gi 125977	-----VOYKAQQGVF-----				
	gi 10728878	-----VOYKAQQGVF-----				
	gi 7290546	-----VASNASDIT-----R-YTRP-				
	gi 1362625	-----VASNASDIT-----R-YTRP-				
40		760	770	780	790	800
	NOV2A	HGNQVSSLLSVRTSESVPDSAPENITYKNISSCEHELSLPPSSPNGIIO				
	gi 12621078	HGNQSSLLSVRTSETVPDSAPENITYKNISSCEHELSLPPSPNGIIO				
	gi 125977	-----SQPSNFRATEIGETATITQTKPHTHSSENIV-----				
	gi 10728878	-----SQPSNFRATEIGETATITQTKPHTHSSENIV-----				
	gi 7290546	-----AAPLIOELASIDEG---LMLFWRSVDVNSRQD-----				


```

gi|1362625| -----AAPLIQELRSTDCG---DMLSWRSDVNSRQD

      810      820      830      840      850
5  NOV2A      KYTFLKRSNG-NEERTTNTTSLTONIKGLKKYTOYITVSASILKCEGV
gi|12621078| KYTFLKRSNS-HEARTTNTTSLTOTIGLKKYTHYVTEVSASILKCEGI
gi|125977|    HYELVNDIVANQAHHKRISNSEAYTIDGLYFDILYYIWLAAARSORGECA
gi|10728878| HYELVNDIVANQAHHKRISNSEAYTIDGLYFDILYYIWLAAARSORGECA
gi|7290546|   RYEVHYCRNGT-REERTNATNETSLTIHYLHEGSGYEVKRVHAISH---GV
10 gi|1362625| RYEVHYCRNGT-REERTNATNETSLTIHYLHEGSGYEVKRVHAISH---GV

      860      870      880      890      900
15 NOV2A      RSRPISILREEDAPDSPPONFSVKOLSGVTVMKLSWOPP--LEPNCIILYY
gi|12621078| RSRPISILREEDAPDSPPONFSVKOLSGVTVMKLSWOPP--LEPNCIILYY
gi|125977|    TUPPIPVRTKQYVPCAPPRNITAIATSTTISLSWLPPEVERSNCRITIVY
gi|10728878| TUPPIPVRTKQYVPCAPPRNITAIATSTTISLSWLPPEVERSNCRITIVY
gi|7290546|   RSEPHSYFOAVFP--KPPONLTLCITVHNLVVLHWQAP--EGST-FSEYV
20 gi|1362625| RSEPHSYFOAVFP--KPPONLTLCITVHNLVVLHWQAP--EGST-FSEYV

      910      920      930      940      950
25 NOV2A      TVYVWR---SSLKITN-VTETSHELSDDLYNVEYSAYVTASTREFGCKT
gi|12621078| TVYVWK---SSLKAIN-ATEASLVLSDDLYNVDYGACVTASTREFGDNA
gi|125977|    KVFVVEVGREDDEATTMTLNMTSTVLDEKRWTEYKTIWVLACTSVGDC-P
gi|10728878| KVFVVEVGREDDEATTMTLNMTSTVLDEKRWTEYKTIWVLACTSVGDC-P
gi|7290546|   VRYRTDA---SPWQRTISGLHENEARKDKHYGERYLVQVNTVS--FGVESF
30 gi|1362625| VRYRTDA---SPWQRTISGLHENEARKDKHYGERYLVQVNTVS--FGVESF

      960      970      980      990      1000
35 NOV2A      RSNITISFCTPEG-PSDPPKDVYYANLSSSSIIILFWTPFS--KPNGIIQYY
gi|12621078| RSSIINFRTPEGEPSDPPNDVHYVNLSSSSIIILFWTPFV--KPNGIIQYY
gi|125977|    RSHPIILRTCEDVPCD-PQDVKATPLNSTSIHVSMKPPLEKDRNGIIRGY
gi|10728878| RSHPIILRTCEDVPCD-PQDVKATPLNSTSIHVSMKPPLEKDRNGIIRGY
gi|7290546|   HPLELNVIMEPQ---PVSNNVPLVDSRNLILEWP---RPDCHVDYF
40 gi|1362625| HPLELNVIMEPQ---PVSNNVPLVDSRNLILEWP---RPDCHVDYF

      1010     1020     1030     1040     1050
45 NOV2A      SVYYRNTSGTFMQNFTLHEVTNDPDNMTVSTIIDKLTFISYYTFWLTAST
gi|12621078| SVYYQNTSGTFVQNFLLQVTKESDNVTVSARIYRLAIFSYTFWLTAST
gi|125977|    -----
gi|10728878| -----
50 gi|7290546| -----
gi|1362625| -----

      1060     1070     1080     1090     1100
55 NOV2A      SVGNKNGKSSDIIIEVYTDQDVEGPFVGNLTYESISSTAINMSWVPPAPNG
gi|12621078| SVGNKNGKSSDIIHVYTDQDIPEGPVGNLTFESISSTAIHMSWPEPSPNG
gi|125977|    -----HIHAQELRDEC
gi|10728878| -----HIHAQELRDEC
gi|7290546|   -----TKKWFTDEEDR
60 gi|1362625| -----TKKWFTDEEDR

      1110     1120     1130     1140     1150
65 NOV2A      EVFYYSGLILQOQTP-RHVRPPLVTYERSYFQMLEKYTDVILKITPSTER
gi|12621078| EVFYYSGLNLQOSPPRHMIPPLVTYENSIDFDLEKYTDVIFKITPSTER
gi|125977|    KGFLNEPFKFDVVD-----TDEFNVTGLCPDTKYSIQVAALTRK
gi|10728878| KGFLNEPFKFDVVD-----TDEFNVTGLCPDTKYSIQVAALTRK
gi|7290546|   MEFKNVITQLEDLSS-----P-SVREPTEDLSFGROYRFEVQASSN-
70 gi|1362625| MEFKNVITQLEDLSS-----P-SVREPTEDLSFGROYRFEVQASSN-

      1160     1170     1180     1190     1200
75 NOV2A      GFSITYTAQYIKKEEDIPETSPINTFNLSSEFSLLSADPPVAPNCAI
gi|12621078| GFSITYTTOYIKKEEDVPDTPPIINTFNLSSEFSLLSADPPVAPNCAI
gi|125977|    GDSRSAAIIVKTPCGVVPVPTVSKIMREPIVSIIEAERERACTYCEI

```

39

gi|12621078| wistotlpgpdpghpennvvaatspfcgniswsepahttgeftnyilndvks
gi|125977| kvttvrik---pedvplnlrahdvsthsmtlswsppirltl-pvnyklisfda
gi|10728878| kvttvrik---pedvplnlrahdvsthsmtlswsppirltl-pvnyklisfda
gi|7290546| tdfadvsysttmrressaptsasyotltappgkvdyfqpdsd
gi|1362625| -----tdfadvsysttmrressaptsasyotltappgkvdyfqpdsd

1610 1620 1630 1640 1650

NOV2A VDNDEFNISFIKSNENKTEIEIKDLEIFTRYSVVITAFNGISAAVVEGK
gi|12621078| VDDDDFNISFLKSNENKTEINNLEVFTRYSVVITAFVGNVSRAYTDGK
gi|125977| MK-----
gi|10728878| MK-----
gi|7290546| VQ-----
gi|1362625| VQ-----

1660 1670 1680 1690 1700

NOV2A SSAEMIVTTLESAPKDPENNMTFOKIPDEVTKFQLTFLPPSQPNENIRVY
gi|12621078| SSAEVIITTTLESVPKDPENNMTFOKIPDEVTKFQLTFLPPSQPNENIRVY
gi|125977| -----VFVDSQGESOT-----QIVPKREIILKH
gi|10728878| -----VFVDSQGESOT-----QIVPKREIILKH
gi|7290546| -----PGEVTFEWS-----LEPAEQHGPIIDYE
gi|1362625| -----PGEVTFEWS-----LEPAEQHGPIIDYE

1710 1720 1730 1740 1750

NOV2A QALVMREDDPTAVQIHNLSIIQKTNTFVIAMLEGKKGCHYNISVYVAVNS
gi|12621078| QALVMREDDPTAVQIHNLSIIQKTNTFVIAMLEGKKGCHYNISVYVAVNS
gi|125977| YMKTHITINELSPFTTYNVNVSAIPS-----DYSYRPETKITVITQMAAPQ
gi|10728878| YMKTHITINELSPFTTYNVNVSAIPS-----DYSYRPETKITVITQMAAPQ
gi|7290546| RITCONADDADVSSYEFPVNAATQG-----KIDGLVPGNHIFRICAQSA
gi|1362625| RITCONADDADVSSYEFPVNAATQG-----KIDGLVPGNHIFRICAQSA

1760 1770 1780 1790 1800

NOV2A ACAGPKVPVMPRIITMDIKAPARPKTKPTITTYDATGKLLVTSATITRMPICY
gi|12621078| ACAGPKVQMRTITMDIKAPARPKSKPIITRDATGKLLVTSATITRMPICY
gi|125977| PMVKPDPFYGVNGEETLVILPOASEEYGPISH-----YYLVVMPEDKSN
gi|10728878| PMVKPDPFYGVNGEETLVILPOASEEYGPISH-----YYLVVMPEDKSN
gi|7290546| LQYGAEREHLOTMPILAPPVPEPSVTPLEVSR-----TSSTIEISFRQCY
gi|1362625| LQYGAEREHLOTMPILAPPVPEPSVTPLEVSR-----TSSTIEISFRQCY

1810 1820 1830 1840 1850

NOV2A YSDDHGPKTKVQVYVITETGAQHDG--NVTIKYDAYENKAR-PYFTNECEP
gi|12621078| YNDDHGPKTRNVQVYVITETGAQQDG--NVTIKYDAYENKAR-PYFTNECEP
gi|125977| LHKIPDQFLITDDLPGRNKPERPN-----APYIAAKPEPRTSIPETFFHLCSG
gi|10728878| LHKIPDQFLITDDLPGRNKPERPN-----APYIAAKPEPRTSIPETFFHLCSG
gi|7290546| FSNAGHVRYSYTIITLAEIVGKASGLEMPSNODVCAYTVWLPYQAIPEPN
gi|1362625| FSNAGHVRYSYTIITLAEIVGKASGLEMPSNODVCAYTVWLPYQAIPEPN

1860 1870 1880 1890 1900

NOV2A NPPTCEGKTKFSGNEEIIYIIGADNACMIPGNEEDKICNGPLRPKKOYLRFK
gi|12621078| NPPTCEGKTKFSGNEEIIYIIGADNACMIPGNEEDKICNGPLRPKKOYLRFK
gi|125977| DDYHNFTRNKLREKRMRYFVRAVVDTPCK--LYTSSPFSEFLSLDMRE
gi|10728878| DDYHNFTRNKLREKRMRYFVRAVVDTPCK--LYTSSPFSEFLSLDMRE
gi|7290546| PFLTSGSRKSSLEAEHFTTIGTANCPKHQAG--YCNGPLRAGTIVYRI
gi|1362625| PFLTSGSRKSSLEAEHFTTIGTANCPKHQAG--YCNGPLRAGTIVYRI

1910 1920 1930 1940 1950

NOV2A RATNIMCGFTDSEYSDPKITLGEGLSERIVEIISVTLCLMSIIEGTAT
gi|12621078| RATNVMCGFTDSEYSDPKITLGEGLSERIVEIISVTLCLMSIIEGTAT
gi|125977| APPGGRPHRPDPNAPPEPEVSNNRNKEPEIENWVPLMUSHFISTALAI
gi|10728878| APPGGRPHRPDPNAPPEPEVSNNRNKEPEIENWVPLMUSHFISTALAI
gi|7290546| RAFTLEDKFTDVIYSSPITTE-----RSDTVIVRAATSAULVAVVLA
gi|1362625| RAFTLEDKFTDVIYSSPITTE-----RSDTVIVRAATSAULVAVVLA

1960 1970 1980 1990 2000

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

FAPARIROKQ-----KEGGTYSQDASINDTKLKLCLITVADLESKDER
FAPVRIROKQ-----KEGGTYSQDASINDTKFKLCLITVADLESKDER
VLCVVKRRRROPCKTPDQAAVTRPELMAADYGAGTPPSQPMVDRRLNFPQTPG
VLCVVKRRRROPCKTPDQAAVTRPELMAADYGAGTPPSQPMVDRRLNFPQTPG
VYCOHFCOLI-----SRASKLAR-----MODELAAPEGYITPN-
VYCOHFCOLI-----SRASKLAR-----MODELAAPEGYITPN-

2010 2020 2030 2040 2050

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

HT-----RPISKKSEFLQVBEELCINNLLKFCSEFSELPKFKLODLSS
HTRLLSYRKSISKPKISKKSEFLQVBEELCINNLLKFCSEFSELPKFKLODLSS
NIS-----HPPIPISSEFANHIERLKSNDNCKFSCEYESTEPG-QOFTW
NIS-----HPPIPISSEFANHIERLKSNDNCKFSCEYESTEPG-QOFTW
RPVHVVKDSESEYRIMSADSDFFRSEFEELKHVGRDOAC
RPVHVVKDSESEYRIMSADSDFFRSEFEELKHVGRDOAC

2060 2070 2080 2090 2100

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

ITADLPWNRANKNRFENIKPYNNNRVKLIADASVPGSDYINASYISGYLCP
ITADLPWNRANKNRFENIKPYNNNRVKLIADVSPGSDYINASYISGYLCP
DNSNLEHNKSKNRMANVTAYDHSRVOLPAVEGVVPGSDYINANYCDGYRKH
DNSNLEHNKSKNRMANVTAYDHSRVOLPAVEGVVPGSDYINANYCDGYRKH
SFANLPCNRPKNRFINILPYDHSRFLQPVDDDDGSDYINANMPCHNSP
SFANLPCNRPKNRFINILPYDHSRFLQPVDDDDGSDYINANMPCHNSP

2110 2120 2130 2140 2150

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

NEFIATQGPLPGTVGDFWRMVEITRAKILVMLTCCFEKGRIRCHQYWPED
NEFIATQGPLPGTVGDFWRMVEITRAKILVMLTCCFEKGRIRCHQYWPED
NAYVATQGLQETFDVDFWRMCWELKTATIVMTRLBERTRIKCDQYWPTR
NAYVATQGLQETFDVDFWRMCWELKTATIVMTRLBERTRIKCDQYWPTR
REFIYVTOGGLPHSTREDFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVC
REFIYVTOGGLPHSTREDFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVC

2160 2170 2180 2190 2200

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

NKPVTVFGDIVILKLMEDVQIDWITIRLKLBERH--GDCMTVROCFETAWP
NKPVTVFGDIVILKLMEDIQIDWITIRLKLBERH--GDCMTVROCFETGWP
G--TETYGCHFTVITETQELATYSIRTFOLCRGFGNDRREIKCLOFTAWP
G--TETYGCHFTVITETQELATYSIRTFOLCRGFGNDRREIKCLOFTAWP
R-VAMFYGDIKVQLIIDTHYHWSISEPMYSRN--CHSRIMRHFFETTPW
R-VAMFYGDIKVQLIIDTHYHWSISEPMYSRN--CHSRIMRHFFETTPW

2210 2220 2230 2240 2250

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

DHGVPEPNSAPLIHFVVLVRASRAHDITPMIVHCSAGVGRGCVFIALDRIT
DHGVPEPNTPLDHFVKLVTRASRAHDITPMIVHCSAGVGRGCVFIALDRIT
DHGVDPHAPFQCTLRRCRAITPPESGPVIVHCSAGVGRGCVFIALDRIT
DHGVDPHAPFQCTLRRCRAITPPESGPVIVHCSAGVGRGCVFIALDRIT
DFGVPEPPQSLVRFVRAFRDVIIGTDMRPIIVHCSAGVGRGCVFIALDRIT
DFGVPEPPLSLVRFVRAFRDVIIGTDMRPIIVHCSAGVGRGCVFIALDRIT

2260 2270 2280 2290 2300

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

QHINHDHFDVDIYGLVAELRSERCMVQNLAQYIFLHCQILDLLS
QHINHDHFDVDIYGLVAELRSERCMVQNLAQYIFLHCQILDLLS
ERMKHEKIDIIYGHVTCLRACNFMVOTEDQYIFIHDAILEAMICG
ERMKHEKIDIIYGHVTCLRACNFMVOTEDQYIFIHDAILEAMICG
QHIRKSDYVDIRGIVFAMRKERYFMVOTEQQYVCIHQCLLAVLEGKEHLL
QHIRKSDYVDIRGIVFAMRKERYFMVOTEQQYVCIHQCLLAVLEGKEHLL

2310 2320 2330 2340 2350

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

-----NKGCSNOPICVNYNSALQKVDLSLAMEGGDVEHEWEEITM
-----NKGCHQPVCEVNYSTLOKMDSLAMEG-DVEHEWEEITM
--VTEVPARNLTHLQKLLITEPGETISGMVEFKKLSNVMDSKPVTA
--VTEVPARNLTHLQKLLITEPGETISGMVEFKKLSNVMDSKPVTA
ADSLSEHANGCEVTVIYLEROPQTKVGLPIRASLAAEKLDDDLMTNK
ADSLSEHANGCEVTVIYLEROPQTKVGLPIRASLAAEKLDDDLMTNK

		2360	2370	2380	2390	2400
5	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	NLPCNKHNNLVHLLPYESSRVYLTPIHGIEGSDYVNASFIDGYRYSAY				
	gi 10728878	NLPCNKHNNLVHLLPYESSRVYLTPIHGIEGSDYVNASFIDGYRYSAY				
	gi 7290546	DEDEDDEEQQQQQQ-----LATEVMPKGSN				
10	gi 1362625	DEDEDDEEQQQQQQ-----LATEVMPKGSN				
		2410	2420	2430	2440	2450
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
15	gi 125977	IAAGPVQDAAEQFWRLWEHNSITIVMLTKLKEMGREKCFQYWPHERSV				
	gi 10728878	IAAGPVQDAAEQFWRLWEHNSITIVMLTKLKEMGREKCFQYWPHERSV				
	gi 7290546	DDEDEEDDDDDDDQPLNNETTATSSASCSSS-----THQVHV				
	gi 1362625	DDEDEEDDDDDDDQPLNNETTATSSASCSSS-----THQVHV				
20		2460	2470	2480	2490	2500
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
25	gi 125977	RYQYYVVDIAEYNMPQYKLREFKVIDARDGSSRTVRQFQFIDWPEQVVP				
	gi 10728878	RYQYYVVDIAEYNMPQYKLREFKVIDARDGSSRTVRQFQFIDWPEQVVP				
	gi 7290546	VLQEAIEKEKQEQERICAGTSHADTESDNTDSDDDDEGDGKVAKCAV				
	gi 1362625	VLQEAIEKEKQEQERICAGTSHADTESDNTDSDDDDEGDGKVAKCAV				
30		2510	2520	2530	2540	2550
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	KSGECFIDFIGQVHKTKEQFGQDGPITVHCSAGVGRSGVPITLSIVLERM				
	gi 10728878	KSGECFIDFIGQVHKTKEQFGQDGPITVHCSAGVGRSGVPITLSIVLERM				
35	gi 7290546	ADEDCWY-----				
	gi 1362625	ADEDCWY-----				
40		2560	2570	2580	2590	
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	QYEGVLDVFQTVRILRSQRPAMVQTEDQYHFCYRAALEYLGSFDNYTN				
	gi 10728878	QYEGVLDVFQTVRILRSQRPAMVQTEDQYHFCYRAALEYLGSFDNYTN				
45	gi 7290546	-----				
	gi 1362625	-----				

Tables 2J-2EE list the domain descriptions from DOMAIN analysis results against NOV2a. This indicates that the NOV2a sequence has properties similar to those of other proteins known to contain this domain.

50

Table 2J. Domain Analysis of NOV2a

gnl|Smart|smart00194, PTPc, Protein tyrosine phosphatase, catalytic domain (SEQ ID NO:93)
 CD-Length = 264 residues, 99.6% aligned
 Score = 318 bits (816), Expect = 2e-87

55

NOV 1: 1983 KFQEEFSELPK-FLQDLSSDADLPWNRAKNRFPNIKPYNNRVKLIADASVPGSDYINA 2041
 +||| +| + ||| | | | | | + + || + ||| |||||
 Sbjct: 1 GLEEEFEKLQRLTPDDLSCVTAILPENRDKNRYKDVLPYDHTRVKL-KPPPGEGSDYINA 59
 NOV 1: 2042 SYISGYLCPNEFIATQGPLPGTVGDFWRMVWETRAKTLVMLTQCFEKGRIRCHQYWPEDN 2101

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 3D.

Table 3D. ClustalW Analysis of NOV3

5	1) NOV3 (SEQ ID NO:12)
	2) ref XP_027243.1 (XM_027243) hypothetical protein XP_027243 (Homo sapiens) (SEQ ID NO:42)
	3) gi 15076843 gb AAK82958.1 AF233450_1 (AF233450) pecanex-like protein 1 (Homo sapiens) (SEQ ID NO:43)
10	4) gi 6650377 gb AAF21809.1 AF096286_1 (AF096286) pecanex 1 (Mus musculus) (SEQ ID NO:44)
	5) gi 13171105 gb AAK13590.1 AF154413_1 (AF154413) pecanex (Takifugu rubripes) (SEQ ID NO:45)
15	6) gi 7290294 gb AAF45755.1 (AE003423) pcx gene product [alt 1] (Drosophila melanogaster) (SEQ ID NO:46)
20	<div> <div> 1020</div> <div> <div>10</div> <div>20</div> <div>30</div> <div>40</div> <div>50</div> <div>60</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> <div>gi 15076843 gb </div> <div>gi 6650377 gb A</div> <div>gi 13171105 gb </div> <div>gi 7290294 gb A</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> <div>MGSQTLQILRQGVWAALSGGWYYPDPHQATFVNALHLYLWLFLLGLPFTLYMALPSTMIIV</div> <div>MGSQTLQILRQGVWASVTGGWYYPDPQNTFVNALHLYIWLFLLCFPFTLYMALQPSMVIV</div> </div> <div> <div>1</div> <div>1</div> <div>60</div> <div>1</div> <div>60</div> <div>1</div> </div> </div>
25	
30	<div> <div>70</div> <div>80</div> <div>90</div> <div>100</div> <div>110</div> <div>120</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> <div>gi 15076843 gb </div> <div>gi 6650377 gb A</div> <div>gi 13171105 gb </div> <div>gi 7290294 gb A</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> <div>AVYCPVIAAVFIVLKMVNYRLHRLDAGEVVDRTANEFTDQR-TKAEQGNCSTRKDSNG</div> <div>GIYCGVIAAMFLLKTVNYRLHHLDEGEVVEHQTRSKSGSGGTGGANDPVTRREDSNG</div> </div> <div> <div>1</div> <div>1</div> <div>119</div> <div>1</div> <div>120</div> <div>1</div> </div>
35	
40	<div> <div>130</div> <div>140</div> <div>150</div> <div>160</div> <div>170</div> <div>180</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> <div>gi 15076843 gb </div> <div>gi 6650377 gb A</div> <div>gi 13171105 gb </div> <div>gi 7290294 gb A</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> <div>PSDPGGGIEMSEFIREATPPVGCSSRNSYAGLDPSNQIGSGSSRLGTAATIKGDTDTAKT</div> <div>LGDPGGGIEMADFIRQETPPVDCSSRNSYVG-----</div> </div> <div> <div>1</div> <div>1</div> <div>179</div> <div>1</div> <div>151</div> <div>1</div> </div>
45	
50	<div> <div>190</div> <div>200</div> <div>210</div> <div>220</div> <div>230</div> <div>240</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> <div>gi 15076843 gb </div> <div>gi 6650377 gb A</div> <div>gi 13171105 gb </div> <div>gi 7290294 gb A</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> <div>SDDISLSLGQSSSLCKEGSEEQDLAADRKLFRLVSNDSFISIQPSLSSCGQDLPRDFSDK</div> <div>-----</div> </div> <div> <div>1</div> <div>1</div> <div>239</div> <div>1</div> <div>151</div> <div>1</div> </div>
55	
60	<div> <div>250</div> <div>260</div> <div>270</div> <div>280</div> <div>290</div> <div>300</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> <div>gi 15076843 gb </div> <div>gi 6650377 gb A</div> <div>gi 13171105 gb </div> <div>gi 7290294 gb A</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> <div>VNLPSHNHHHVDQSLSSACDTEVASLVPLHSHSYRKDHRPRGVPTSSSAVAFPDTSLN</div> <div>-----</div> </div> <div> <div>1</div> <div>1</div> <div>299</div> <div>1</div> <div>151</div> <div>1</div> </div>
65	
	<div> <div>310</div> <div>320</div> <div>330</div> <div>340</div> <div>350</div> <div>360</div> </div> <div> <div>NOV3</div> <div>ref XP_027243.1</div> </div> <div> <div>..... </div> <div>----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- </div> </div> <div> <div>1</div> <div>1</div> </div>

56



NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

5

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

10

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

15

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

20

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

25

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

30

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

35

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

40

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

45

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

50

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

55

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

60

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

65

NOV3
ref|XP_027243.1
gi|15076843|gb|
gi|6650377|gb|A
gi|13171105|gb|
gi|7290294|gb|A

70

		1270	1280	1290	1300	1310	1320				
5	NOV3	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 412			
	ref XP_027243.1	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 1			
	gi 15076843 gb	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 1307			
	gi 6650377 gb A	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 412			
	gi 13171105 gb	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 716			
	gi 7290294 gb A	DPLPEKLRNSV	SERLQSDLV	VVCIVL	GVLYFAIHVSTV	TVLQPAK	KYVLYL	LVGVGVMT 1			
10		1330	1340	1350	1360	1370	1380				
	NOV3	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 472			
	ref XP_027243.1	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 1			
	gi 15076843 gb	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 1367			
15	gi 6650377 gb A	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 472			
	gi 13171105 gb	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 776			
	gi 7290294 gb A	HYVLPQVRKQLP	WHCFSPHLLK	TLLEYNQY	EVNRDAAT	MMWFEK	LHVWLLF	VEKNIIYPLIV 1			
20		1390	1400	1410	1420	1430	1440				
	NOV3	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 532			
	ref XP_027243.1	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 1			
	gi 15076843 gb	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 1427			
25	gi 6650377 gb A	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 532			
	gi 13171105 gb	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 836			
	gi 7290294 gb A	LNELSSAETIAS	PKRLNTELC	ALMITV	AGLKLLRSS	SSPTYOY	ITVIT	FTVLFFKFDYE 1			
30		1450	1460	1470	1480	1490	1500				
	NOV3	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 592			
	ref XP_027243.1	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 1			
	gi 15076843 gb	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 1487			
	gi 6650377 gb A	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 592			
35	gi 13171105 gb	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 896			
	gi 7290294 gb A	AFSETMLLDLE	FMSTILFN	KLWELLY	KLOFVYTY	IAPWQIT	WGSAFH	AFAPFAVPHSAML 1			
40		1510	1520	1530	1540	1550	1560				
	NOV3	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	652		
	ref XP_027243.1	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	1		
	gi 15076843 gb	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	1547		
	gi 6650377 gb A	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	652		
	gi 13171105 gb	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	956		
45	gi 7290294 gb A	FVQAIVSAFF	STPLNPF	LGSATF	ITSYVR	PKFWERD	YNTKR	VDHSNTRLASOLDRNP	47		
		1570	1580	1590	1600	1610	1620				
	NOV3	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 712		
	ref XP_027243.1	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 1		
50	gi 15076843 gb	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 1607		
	gi 6650377 gb A	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 712		
	gi 13171105 gb	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 1016		
	gi 7290294 gb A	DDNNLNSIF	YEHLTR	SLQHS	LCCDLL	GRWGN	YSTGDC	FILASDYL	NALVHLIEIGNGLV 97		
55		1630	1640	1650	1660	1670	1680				
	NOV3	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 772	
	ref XP_027243.1	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 1	
	gi 15076843 gb	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 1667	
60	gi 6650377 gb A	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 772	
	gi 13171105 gb	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 1076	
	gi 7290294 gb A	TFQLRGL	EFRGTY	CQORE	VEAITE	GVEE	DEGFC	CCPEPGH	IPHLSFNAAFS	QORWLAWEV 105	
65		1690	1700	1710	1720	1730	1740				
	NOV3	VTKYILE	GYSITD	NSAAS	MLQV	FLRR	YLT	TYVVK	GIIYV	VITSSKLEEWLAN	ETMQEGL 832
	ref XP_027243.1	VTKYILE	GYSITD	NSAAS	MLQV	FLRR	YLT	TYVVK	GIIYV	VITSSKLEEWLAN	ETMQEGL 5
	gi 15076843 gb	VTKYILE	GYSITD	NSAAS	MLQV	FLRR	YLT	TYVVK	GIIYV	VITSSKLEEWLAN	ETMQEGL 1727
	gi 6650377 gb A	VTKYILE	GYSITD	NSAAS	MLQV	FLRR	YLT	TYVVK	GIIYV	VITSSKLEEWLAN	ETMQEGL 832
70	gi 13171105 gb	VTKYILE	GYSITD	NSAAS	MLQV	FLRR	YLT	TYVVK	GIIYV	VITSSKLEEWLAN	ETMQEGL 1136

	gi 6650377 gb A	GFVPCRRSSTSQISLRNLPSSIQSRLSMVNQMEARS	CGG	CCVOHCLPSSSSSSSQSIPAC	1308		
	gi 13171105 gb	GLEPCRRSSTSQISLRNLPSTISLRIGST	SDPAG	SSSSSSSHSIFPC	1558		
	gi 7290294 gb A	--G-----VKKLEAFPFFFTLLFGKG	RRYAC	HC--SYKDVIIDGSM	485		
5		2230	2240	2250	2260	2270	2280
	NOV3	KHHTLVGFLATEGGQSSATDAQ	PGNTLSPANNSHS	RKA	EVIYRVOIVDFSOIL	1361	
	ref XP_027243.1	KHHTLVGFLATEGGQSSATDAQ	PGNTLSPANNSHS	RKA	EVIYRVOIVDFSOIL	534	
	gi 15076843 gb	KHHTLVGFLATEGGQSSATDAQ	PGNTLSPANNSHS	RKA	EVIYRVOIVDFSOIL	2256	
10	gi 6650377 gb A	KHHTLVAFELGEGGQSSATEAC	PGNTLSPANNISHA	RKGE	EVIYRVOIVDFSOIL	1361	
	gi 13171105 gb	KHHTLVGFLGNDGLCSITVTCPLSQHHHPHHHPQHNP	PHATVRRD	DISYRVOIVDFSOIL	1618		
	gi 7290294 gb A	EDFKFVAFINEADTVLAVASC	G	RDPIVAQFAELTSQCKCL	526		
15		2290	2300	2310	2320	2330	2340
	NOV3	EGINLSKRKELQWPDEGIRLKAGRNSWKDW	SPOEGMEGHVIHRWVPCSRDPGTRSHIDKA	1421			
	ref XP_027243.1	EGINLSKRKELQWPDEGIRLKAGRNSWKDW	SPOEGMEGHVIHRWVPCSRDPGTRSHIDKA	594			
	gi 15076843 gb	EGINLSKRKELQWPDEGIRLKAGRNSWKDW	SPOEGMEGHVIHRWVPCSRDPGTRSHIDKA	2316			
	gi 6650377 gb A	EGINLSKRKELQWPDEGIRLKAGRNSWKDW	SPOEGMEGHVIHRWVPCSRDPGTRSHIDNT	1421			
20	gi 13171105 gb	EGINLSKRKELQWPDEGIRLKAGRNSWKDW	SPOEGMEGHVIHRWVPCSRDPGTRSHIDNT	1678			
	gi 7290294 gb A	G-----RGOIEDP-----A-----RREDWTKKLCOP-----LPQVR-----	552				
25		2350	2360				
	NOV3	VLLVQIDDKYVTVIETGVLELGAEV	1446				
	ref XP_027243.1	VLLVQIDDKYVTVIETGVLELGAEV	619				
	gi 15076843 gb	VLLVQIDDKYVTVIETGVLELGAEV	2341				
	gi 6650377 gb A	VLLVQIDDKYVTVIETGVLELGAEV	1446				
	gi 13171105 gb	VLLVQIDDKYVTVIETGVLELGAEV	1703				
30	gi 7290294 gb A	-----	552				

Pecanex gene was originally discovered in *Drosophila*, encoding a large, membrane-spanning protein. The mouse homolog was recently reported. In the absence of maternal expression of the pecanex gene, the embryo develops severe hyperneuralization similar to that characteristic of Notch mutant embryos. Early gastrula embryos, lacking both maternally and zygotically expressed activity of the neurogenic pecanex locus, are shown to contain a greater than wild-type number of stably determined neural precursor cells which can differentiate into neurons in culture. Therefore it is anticipated that this novel human pecanex will be involved in neuronal differentiation, maintenance of neuronal precursors and neurological diseases.

The disclosed NOV3 nucleic acid of the invention encoding a Human homolog of the *Drosophila* pecanex protein includes the nucleic acid whose sequence is provided in Table 3A or a fragment thereof. The invention also includes a mutant or variant nucleic acid any of whose bases may be changed from the corresponding base shown in Table 3A while still encoding a protein that maintains its Human homolog of the *Drosophila* pecanex activities and physiological functions, or a fragment of such a nucleic acid. The invention further includes nucleic acids whose sequences are complementary to those just described, including nucleic acid fragments that are complementary to any of the nucleic acids just described. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications. Such modifications include, by way

gi 13653970 ref XP_009546.3 (XM_009546)	serine/threonine kinase 15 [Homo sapiens]	403	369/403 (91%)	381/403 (93%)	0.0
gi 4507275 ref NP_03591.1 (NM_003600)	serine/threonine kinase 15; Serine/threonine protein kinase 15 (Homo sapiens)	403	369/403 (91%)	380/403 (93%)	0.0
gi 7446411 pir JC5974	aurora-related kinase 1 (EC 2.7.-.-) - human	403	367/403 (91%)	379/403 (93%)	0.0
gi 4507279 ref NP_03149.1 (NM_003158)	serine/threonine kinase 6; Serine/threonine protein kinase-6; serine/threonine kinase 6 (aurora/IPL1-like) [Homo sapiens]	402	342/403 (84%)	360/403 (88%)	0.0

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 4D.

5

Table 4D. ClustalW Analysis of NOV4

- 1) NOV4 (SEQ ID NO:14)
- 2) gi|12654873|gb|AAH01280.1|AAH01280 (BC001280) serine/threonine kinase 15 (Homo sapiens) (SEQ ID NO:47)
- 3) gi|13653970|ref|XP_009546.3| (XM_009546) serine/threonine kinase 15 (Homo sapiens) (SEQ ID NO:48)
- 4) gi|4507275|ref|NP_03591.1| (NM_003600) serine/threonine kinase 15; Serine/threonine protein kinase 15 (Homo sapiens) (SEQ ID NO:49)
- 5) gi|7446411|pir|JC5974 aurora-related kinase 1 (EC 2.7.-.-) - human (SEQ ID NO:50)
- 6) gi|4507279|ref|NP_03149.1| (NM_003158) serine/threonine kinase 6; Serine/threonine protein kinase-6; serine/threonine kinase 6 (aurora/IPL1-like) (Homo sapiens) (SEQ ID NO:51)

20

25

30

35

		10	20	30	40	50
NOV4	MD	SE	ENCIS	GPV	SA	KTPVGGP
gi 12654873	MD	SK	ENCIS	GPV	KAT	APVGGP
gi 13653970	MD	SK	ENCIS	GPV	KAT	APVGGP
gi 4507275	MD	SK	ENCIS	GPV	KAT	APVGGP
gi 7446411	MD	SK	ENCIS	GPV	KAT	APVGGP
gi 4507279	MD	SK	ENCIS	GPV	KAT	APVGGP
		60	70	80	90	100
NOV4	SN	SS	Q	R	V	P
gi 12654873	SN	SS	Q	R	V	P
gi 13653970	SN	SS	Q	R	V	P
gi 4507275	SN	SS	Q	R	V	P
gi 7446411	SN	SS	Q	R	V	P

gi|4507279| SSSQSRVPLQAQKLVSSSHKVPQNKKQLQKQLQATSVPHPVSRPLNNTQSKSK

.....110.....120.....130.....140.....150
5 NOV4 SPLSAPENNPEEELASKQKNNEESKKRQWALEDFEIGRPPLGKGKFGNVYLA
gi|12654873| PLPSAPENNPEEELASKQKNNEESKKRQWALEDFEIGRPPLGKGKFGNVYLA
gi|13653970| PLPSAPENNPEEELASKQKNNEESKKRQWALEDFEIGRPPLGKGKFGNVYLA
gi|4507275| PLPSAPENNPEEELASKQKNNEESKKRQWALEDFEIGRPPLGKGKFGNVYLA
10 gi|7446411| PLPSAPENNPEEELASKQKNNEESKKRQWALEDFEIGRPPLGKGKFGNVYLA
gi|4507279| PLPSHLYITLRRNHONAK-MONCKEAVALEDFEIGRPPLGKGKFGNVYLA

.....160.....170.....180.....190.....200
15 NOV4 REKQSKFILALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF
gi|12654873| REKQSKFILALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF
gi|13653970| REKQSKFILALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF
gi|4507275| REKQSKFILALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF
20 gi|7446411| REKQSGTILALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF
gi|4507279| REKQSKFTLALKVLFFKAQLEKAGVEHQLRREVEIQSHLRHPNITRLRYGYF

.....210.....220.....230.....240.....250
NOV4 HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK
25 gi|12654873| HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK
gi|13653970| HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK
gi|4507275| HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK
gi|7446411| HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK
gi|4507279| HDATRYYLILEYAPLGTVYRELQKLSKFDEORTATYITELANALSYCHSK

.....260.....270.....280.....290.....300
30 NOV4 TVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM
gi|12654873| RVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM
35 gi|13653970| RVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM
gi|4507275| RVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM
gi|7446411| RVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM
gi|4507279| RVIHRDIKPENLLGSAGELKIADFGWSVHAPSSRRRTTLCGTLDYLPPPM

.....310.....320.....330.....340.....350
40 NOV4 IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD
gi|12654873| IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD
gi|13653970| IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD
45 gi|4507275| IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD
gi|7446411| IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD
gi|4507279| IEGRMHDEKVDLWSGLVLCYEFVLVGKPPFEANTYQETKYKRISRVEFTFPD

.....360.....370.....380.....390.....400
50 NOV4 FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS
gi|12654873| FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS
gi|13653970| FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS
55 gi|4507275| FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS
gi|7446411| FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS
gi|4507279| FVTLEGARDLISRLKKHNPSQRPMLEVLHEHPWITANSSKPSNCNQNKESAS

.....
60 NOV4 KYS
gi|12654873| KQS
gi|13653970| KQS
gi|4507275| KQS
gi|7446411| KQS
65 gi|4507279| KQS

Table 5B. NOV5 protein sequence (SEQ ID NO:16)

MGQSQGDGHGPRRGKKDEKDKKNKYEPVPTRVAEKEEKTGQDVASKLPLVTLHTQCRLKLLKLERIKDYLLM
VEEFIRNQEIQIKLLEEKQEEGRSKVDDLRGTPMSVGNLEEIIDDNHAIVSTSVGSEHYDSIIISFVEKDILLEPGC
SILLRHKVHAVIGVLMDDTGPLVTMMKVEKAPQETVYNTGGLDNQIQEIKESMELPLPHPEYYEEMGTKPPKGV
ILCGPPGTGKTLAKAVANQTSATFLRVVGSELIQKYLGDGPKLVRFQVAEEHAPSIMFTDEIEAIGTKRYD
SNSGGEREIQQTMLELLELNQLGGFDSREDVKVIMATKQVETLDPVLRPGRIDKKIEFHLPEKTKKHIFQIH
TSRMTLANDVTLLDLIMAKDDFSGADIKAICTEAGLMALREHRMKATNEDFKKSIESVLYKKHGEIPEGLYL

The full amino acid sequence of the protein of the invention was found to have 383 of 442 amino acid residues (86%) identical to, and 405 of 442 amino acid residues (91%) similar to, the 440 amino acid residue ptnr:SWISSPROT-ACC:P49014 protein from *Mus musculus* (Mouse), and *Rattus norvegicus* (Rat) (26S Protease Regulatory Subunit 4 (P26S4) (E = 1.7e⁻²⁰⁰).

NOV5 also has homology to the amino acid sequences shown in the BLASTP data listed in Table 5C.

Table 5C. BLAST results for NOV5

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi 4506207 ref NP_02793.1 (NM_002802)	proteasome (prosome, macropain) 26S subunit, ATPase, 1; Proteasome 26S subunit, ATPase, 1 [Homo sapiens]	440	382/442 (86%)	405/442 (91%)	0.0
gi 6679501 ref NP_032973.1 (NM_008947)	protease (prosome, macropain) 26S subunit, ATPase 1 [Mus musculus]	440	383/442 (86%)	405/442 (90%)	0.0
gi 345717 pir A44468	26S proteasome regulatory chain 4 [validated] - human	440	381/442 (86%)	404/442 (91%)	0.0
gi 16741033 gb AAH16368.1 AAH16368 (BC016368)	protease (prosome, macropain) 26S subunit, ATPase 1 [Homo sapiens]	440	382/442 (86%)	404/442 (90%)	0.0
gi 2492516 sp Q90732 PRS4_CHICK	26S PROTEASE REGULATORY SUBUNIT 4 (P26S4)	440	378/442 (85%)	402/442 (90%)	0.0

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 5D.

Table 5D ClustalW Analysis of NOV5

- 1) NOV5 (SEQ ID NO:16)
- 2) gi|4506207|ref|NP_002793.1| (NM_002802) proteasome (prosome, macropain) 26S subunit, ATPase, 1; Proteasome 26S subunit, ATPase, 1 [Homo sapiens] (SEQ ID NO:52)
- 3) gi|6679501|ref|NP_032973.1| (NM_008947) protease (prosome, macropain) 26S subunit, ATPase 1 [Mus musculus] (SEQ ID NO:53)
- 4) gi|345717|pir||A44468 26S proteasome regulatory chain 4 [validated] - human (SEQ ID NO:54)

NOV5	RKIEFPLPDEKTKKRIFQIH	TSRMTLAN	DVTLDDLIMAKDD	ESGADIKAI
gi 4506207	RKIEFPLPDEKTKKRIFQIH	TSRMTLADDVTLDDLIMAKDD	LSGADIKAI	
gi 6679501	RKIEFPLPDEKTKKRIFQIH	TSRMTLADDVTLDDLIMAKDD	LSGADIKAI	
gi 345717	RKIEFPLPDEKTKKRIFQIH	TSRMTLADDVTLDDLIMAKDD	LSGADIKAI	
5 gi 16741033	RKIEFPLPDEKTKKRIFQIH	TSRMTLADDVTLDDLIMAKDD	LSGADIKAI	
gi 2492516	RKIEFPLPDEKTKKRIFQIH	TSRMTLADDVTLD	ELIMAKDDLSGADIKAI	

	410	420	430	440
NOV5	CTEAGLMALRERRMKVTNEDFKKSTESVLYKKHEGTP	PEGLYL		
gi 4506207	CTEAGLMALRERRMKVTNEDFKKSKENVLYKKQEGT	PEGLYL		
gi 6679501	CTEAGLMALRERRMKVTNEDFKKSKENVLYKKQEGT	PEGLYL		
gi 345717	CTEAGLMALRERRMKVTNEDFKKSKENVLYKKQEGT	PEGLYL		
gi 16741033	CTEAGLMALRERRMKVTNEDFKKSKENVLYKKQEGT	PEGLYL		
15 gi 2492516	CTEAGLMALRERRMKVTNEDFKKSKENVLYKKHEGTP	PEGLYL		

Tables 5E-F list the domain description from DOMAIN analysis results against NOV5.

This indicates that the NOV5 sequence has properties similar to those of other proteins known to contain this domain.

Table 5E. Domain Analysis of NOV5

gnl|Pfam|pfam00004, AAA, ATPase family associated with various cellular activities (AAA). AAA family proteins often perform chaperone-like functions that assist in the assembly, operation, or disassembly of protein complexes (SEQ ID NO:101)
CD-Length = 186 residues, 100.0% aligned
Score = 190 bits (483), Expect = 1e-49

NOV 4:	221	GVILCGPPGTGKTL	LAKAVANQTSATFLRVV	GSELIQKYLGDGPKLVRQVFQVAEEHAPS	280
Sbjct:	1	GILLYGPPGTGKTL	LAKAVAKELGVFFIEISGSELLSKYVGESEKLV	RALFSLARKSAPC	60
NOV 4:	281	IMFTDEIEAIGTKR	YDSNSGGEREIQOTM	LELELLNQLCGFDSREDVKVIMATKQVETLD	340
Sbjct:	61	IIFIDEIDALAPKRG	DVGTGDVSS---RVVNQLL	TEMDFEKLNSNVIVIGATNRPDLLD	116
NOV 4:	341	PVLIRPGRIDKKIE	FHLPDEKTKKRIFQIH	TSRMTLANDVTLDDLIMAKDD	ESGADIKAI 400
Sbjct:	117	PALLRPGRFDRRIE	VPLPDEEERLEILKIHLKKKPLEKD	VLDLDEIARRTPGFGSADLAAL	176
NOV 4:	401	CTEAGLMALR	410		
Sbjct:	177	CREAALRAIR	186		

Table 5F. Domain Analysis of NOV5

gnl|Smart|smart00382, AAA, ATPases associated with a variety of cellular activities; AAA. This profile/alignment only detects a fraction of this vast family. The poorly conserved N-terminal helix is missing from the alignment. (SEQ ID NO:102)
CD-Length = 151 residues, 100.0% aligned
Score = 61.6 bits (148), Expect = 9e-11

NOV 4:	218	PPKGVILCGPPGTGKTL	LAKAVANQTSATFLRVV-----	GSELIQK	258
Sbjct:	1	PGEVVLIVPPGSGKTT	LARALARELGPDGGGVYIDGEDL	REEALLQLRLVLVGEDK	60

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 6D.

Table 6D Clustal W Sequence Alignment

5	1) NOV6 (SEQ ID NO:18)
	2) gi 3077703 dbj BAA25784.1 (AB004816) mitsugumin29 [Oryctolagus cuniculus] (SEQ ID NO:57)
	3) gi 6678874 ref NP_032622.1 (NM_008596) mitsugumin 29 [Mus musculus] (SEQ ID NO:58)
10	4) gi 12836843 dbj BAB23831.1 (AK005132) putative [Mus musculus] (SEQ ID NO:59)
	5) gi 1351168 sp P20488 SYPH_BOVIN SYNAPTOPHYSIN (MAJOR SYNAPTIC VESICLE PROTEIN P38) (SEQ ID NO:60)
	6) gi 2134413 pir I50720 synaptophysin IIa - chicken (SEQ ID NO:61)
15	
	<div> <div> 1020304050 </div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> MSSTESAGRTADKSPRQVDRLLVGLRWRRLEELGFIKVLQWLFAIFAF MSSTESPSRAADKSPRQVDRLLVGLRWRRLEELGFIKVLQWLFAIFAF MSSTESPGRTSDKSPRQVDRLLVGLRWRRLEELGFIKVLQWLFAIFAF -----MDPVSOVASAGTFRAKLEPLAFLRALELLFAMFAF -----MDVVMQLVAGCGFRVWKEPLGFIKVLQWLFAIFAF -----MCVVLFAPLFAIFAF </div> </div>
20	
25	
	<div> <div>60708090100</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> GSCGSYSGETGAMVRONNEAKDVSSIIVAFGYPPFRLHRTQVEMELQDEES GSCGSYSGETGAMVRONNEAKDVSSIIVAFGYPPFRLHRTQVEMELQDDDS GSCGSYSGETGAMVRONNEAKDVSSIIVAFGYPPFRLHRTQVEMELQDDDS ATCGGYSGLRLSVDCVKNKTESNLSIDTAFAYPPFRLQCVTFEVPFQSGK- ATCGGYSGLRLSVDCVKNKTESNLSIDTAFAYPPFRLHRTQVEMELQDDDS ATCGGYSGLRLSVDCVKNKTESNLSIDTAFAYPPFRLHRTQVEMELQDDDS </div> </div>
30	
35	
	<div> <div>110120130140150</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> SSKTMHLMGDFSAPAEFFVTLGTFSSFYTMAALVVIIRFHKLTYTENKRRFP SSKTMHLMGDFSAPAEFFVTLGTFSSFYTMAALVVIIRFHKLTYTENKRRFP SSKTMHLMGDFSAPAEFFVTLGTFSSFYTMAALVVIIRFHKLTYTENKRRFP EQCKIALVGDSSSAEAEFFVTVAAPFALYSIAATVVVIFFOHKYRENNKGP DPKRIFLVGNYSSEAEFFVTVAAPFALYSIAATVVVIFFOHKYRENNKGP RRFUSLIGDFSSEAEFFVTVAAPFALYSIAATVVVIFFOHKYRENNKGP </div> </div>
40	
45	
	<div> <div>160170180190200</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LIDFIVTVVFSRLWLVCSSAAKGLSDVKIATDPEVLLMSACKQPSNK MLDPLATAVFAFMWLVSAAKGLSDVKIATDPEVLLMSACKQPSNK LIDFIVTVVFSRLWLVCSSAAKGLSDVKIATDPEVLLMSACKQPSNK </div> </div>
50	
55	
	<div> <div>210220230240250</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGCGGCGDQDQ CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGCGGCGDQDQ CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGCGGCGDQDQ QMAVHSPVMSSLNTSVVFGFLNFIWAGNCFVFKETPNHGCGGCGDQDQ CKELRDPVTSGLNTSVVFGFLNFIWAGNCFVFKETPNHGCGGCGDQDQ CLPVRSPVMSSLNTSVVFGFLNFIWAGNCFVFKETPNHGCGGCGDQDQ </div> </div>
60	
65	
	<div> <div>260270280290300</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 </div> <div> DQDQ-----GQGF--SQESAEEQG----- --DQ-----GQGF--SQESAEEQG----- --DQ-----GQGF--SQESAEEQG----- MEKH-----SSSYNG--RYN-QESYGSSEGY--QQAN-----L PEKCPAPGDAYGQAGYGQGF--GYGPOESYGPQGYQPDYGPASSGGGY </div> </div>

Table 6E. Domain Analysis of NOV6

gnl|Pfam|pfam01284, Synaptophysin, Synaptophysin / synaptoporin. (SEQ
ID NO:103)
CD-Length = 298 residues, 70.8% aligned
Score = 244 bits (622), Expect = 6e-66

20	NOV 5:	29	RRLEELPGFIKVLQWLFAIFAFGSCGSYSGETGAMVRCNNEAKDVSSIIIVAFGYPFRLHR	88
	Sbjct:	3	MVIFAPLPGFVKVLQWVFAIFAFATCGGYSSELQLSVDCANKTESDLNIDIAFAYPFRLHE	62
25	NOV 5:	89	IQYEMPLCDEESSKTMHLMGDFSAPAEFFVTLGIFSFFYTMAALVIYLRFHNLTYTENKR	148
	Sbjct:	63	VTFEAPTC-EGDEKQNIALVGDSSSAEEFVTVAVFALYSLAALATYIFQNKYRENNK	121
30	NOV 5:	149	FPLVDFCVTVSFTFFWLVAANAAGKGLTDVKGATRPSSSLTAAMSVCHGEEAVCSAGATPS	208
	Sbjct:	122	GPLIDFIATAVFAFLWLVGSSAWAKGLSDVKMATDPEEIIKGMHACHQPKNCKELHDPV	181
30	NOV 5:	209	MGLANISVLFPGFINFFLWAGNCWFVFKETPPWH	240
	Sbjct:	182	MSGLNTSVVFGFLNFIWLWAGNIWFVKETGWA	213

79

gi|16303264|MRPPPALALAGLCCLALPAAASVYFGLTGRE-----VLTPFFGL
gi|3915306|MALLRALLG--LACTPRPSARYFGLTENE-----ALTILP-L
gi|15082261| MLDGSLARWLAAAFGLTLAALRPSARYFGLTGE-----PLTILP-L
gi|139748|MRILTFLLGLKTWVLAFFSSLENTIAVNNSGKWWGIVAVASAG
gi|3024851|
.....60.....70.....80.....90.....100
NOV7GTAAAPAGGGAHLKQCDLLKLSRROKOLCRREPGLAETLRDAAHGLGLEC
gi|16303264|GTAAAPAGGGAHLKQCDLLKLSRROKOLCRREPGLAETLRDAAHGLGLEC
gi|3915306|TSEMEDAAVKAHYKVCRLKLEKKORMCRDPGCAETLMEATISVSALEC
gi|15082261|TLEP-EAAQAQAHYKACDRLKLERKORMCRDPGCAETLMEAVSVSALEC
gi|139748|NVLPGSDARPVPLVLDPSCQLLSRQKRITRONPGCLQSTITRGLHSAIREC
gi|3024851|
.....110.....120.....130.....140.....150
NOV7QOQFRHERWNCGLEC---RMG---LLKRGFKETAFLYAVSSAALTHILARAC
gi|16303264|QOQFRHERWNCGLEC---RTG---LLKRGFKETAFLYAVSSAALTHILARAC
gi|3915306|QOQFRHERWNCITLEGYRAS---LLKRGFKETAFLYAISSAGLTHAMAKAC
gi|15082261|QOQFRFERWNCITLEGYRAS---LLKRGFKETAFLYAISSAGLTHALAKAC
gi|139748|KWHFRNRWRNGPTGTGNCQVFGKILNRGCRETAFVFAITSAQVTHSVARS
gi|3024851|
.....160.....170.....180.....190.....200
NOV7SAGRMERCTCDDSPGLESRCAWQWCVCGDNLKYSTKFLSNFLGSKRGNKD
gi|16303264|SAGRMERCTCDDSPGLESRCAWQWCVCGDNLKYSTKFLSNFLGSKRGNKD
gi|3915306|SAGRMERCTCDEAPDENRDAWQWCCGDNLKYSNKFVKEFLC-RKPNKD
gi|15082261|SAGRMERCTCDEAPDENRDAWQWCCGDNLKYSKFKVKEFLC-RRSSKD
gi|139748|SEGSIESCSCDYRRRGPGGPDWHWCCSDNTEGRFIGREFVDSSEGRKE
gi|3024851|
.....210.....220.....230.....240.....250
NOV7LRARADAHNTHVGITAVKSGRLRTTCKCHGVSGSCAVRTCWKQLSPFRETC
gi|16303264|LRARADAHNTHVGITAVKSGRLRTTCKCHGVSGSCAVRTCWKQLSPFRETC
gi|3915306|LRARVDFHNLVGMKVTKAGVETTCCKCHGVSGSCVVRTCWRQLSPFHEIC
gi|15082261|LRARVDFHNLVGMKVTKAGVETTCCKCHGVSGSCVVRTCWRQLSPFHEIC
gi|139748|LKYLVNLHNNQACRLTLVTEMROECKCHCMSSGSCSLRTCWMRLDPFRSVG
gi|3024851|SGSCAVRTCWKQLSPFRETC
.....260.....270.....280.....290.....300
NOV7QVLKURYDSAVKVSSATNEALGRLELWAPAR---QGS LTKGLAPRSGDLV
gi|16303264|QVLKURYDSAVKVSSATNEALGRLELWAPAR---QGS LTKGLAPRSGDLV
gi|3915306|QQLKQKYVETSLKVGSTTNEATGE-GDISPPK---KSIPGHSDQIPRTIDLV
gi|15082261|KHLKHVYETALKVGSITNEAAGEAGAISSPRGRASGAGSDPLRPIPELV
gi|139748|DALKDRFEDGASKVIVYNNNGSNRWGSRSDPPH---LAPENTHALPSSODLV
gi|3024851|QVLKURYDSAVKVSSATNEALGRLELWAPAR---QGS LTKGLAPRSGDLV
.....310.....320.....330.....340.....350
NOV7YMEDSPSFCRPSKYS---PGTAGRVCSRE---ASCSS LCCGRGYDTQSRIL
gi|16303264|YMEDSPSFCRPSKYS---PGTAGRVCSRE---ASCSS LCCGRGYDTQSRIL
gi|3915306|YICDPSFCLMSRYS---PGTSGRKQYCI---KNQDS LCCGRGHNTQSRV
gi|15082261|HLDSPSFCLAGRFSS---PGTAGRCHRE---KNQES LCCGRGHNTQSRV
gi|139748|YFEKSPNFCSPSKNKGTPTGTGRICNSTSLGLDGCCELLCCGRGYRSLAEK
gi|3024851|YMEDSPSFCRPSKYS---PGTAGRVCSRE---ASCSS LCCGRGYDTQSRIL
.....360.....370.....380
NOV7VAFSCHCOVWCCVYECQOCVCELVYTCKH
gi|16303264|VAFSCHCOVWCCVYECQOCVCELVYTCKH
gi|3915306|VTRPCCQVRWCCVYECQOCVCELVYTCKD
gi|15082261|VTRPCCQVRWCCVYECQOCVCELVYTCKG
gi|139748|VTERCHCTFNVCCQVETQLNCTSSQIVHCEL
gi|3024851|VAFSCHCOV

VEAVSMSALECQFQFRFERWNCTLEGRYRASLLKRGFKETAFLYAISSAGLTHALAKACSAGRMRCTCDE
APDLENREGWKWGGCSEDI EFGGMVSREFADARENRPDARSAMNRHNNEAGRQVIKAGVETTCCKHGVSGS
CTVRTCWRQLAPFHEVGKHLKHYESALKVGGSTTNEAAGEAGAI SPPRGRASGAGSDPLPRTPELVHLLD
SPSFCLAGRFSPGTAGRRCHREKNCEI CCGRGHNTQSRVVTTRPCQCQVRWCCYVECRQCTQREEVYTCKG

The disclosed NOV8 amino acid sequence has 270 of 354 amino acid residues (76%) identical to, and 310 of 354 amino acid residues (87%) similar to, the 354 amino acid residue ptmr:SWISSPROT-ACC:O42280 protein from *Gallus gallus* (Chicken) (WNT-14 Protein Precursor (1.2e⁻¹⁵¹).

NOV8 is expressed in at least brain. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Literature sources, and/or RACE sources.

In addition, the sequence is predicted to be expressed in brain because of the expression pattern of (GENBANK-ID: gb:GENBANK-ID:AF031168|acc:AF031168.1) a closely related [*Gallus gallus* Wnt-14 protein (Wnt-14) mRNA, complete cds].

NOV8 also has homology to the amino acid sequence shown in the BLASTP data listed in Table 8C.

Table 8C. BLAST results for NOV8					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi 15082261 ref NP_003386.1 (NM_003395)	wingless-type MMTV integration site family, member 14 [Homo sapiens]	365	306/340 (90%)	321/340 (94%)	e-167
gi 3915306 sp O42280 WN14 CHICK	WNT-14 PROTEIN PRECURSOR	354	270/357 (75%)	310/357 (86%)	e-142
gi 16303264 dbj BAB70499.1 (AB063483)	WNT14B [Homo sapiens]	357	193/339 (56%)	244/339 (71%)	e-100
gi 7106447 ref NP_033548.1 (NM_009522)	wingless-related MMTV integration site 3A [Mus musculus]	352	141/311 (45%)	179/311 (57%)	2e-62
gi 5821261 dbj BAA83743.1 (AB024080)	Wnt-3a [Gallus gallus]	376	139/311 (44%)	179/311 (56%)	3e-62

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 8D.

Table 8D. Information for the ClustalW proteins

- 1) NOV8 (SEQ ID NO:22)
- 2) gi|15082261|ref|NP_003386.1| (NM_003395) wingless-type MMTV integration site family, member 14 [Homo sapiens] (SEQ ID NO:64)
- 3) gi|3915306|sp|O42280|WN14_CHICK WNT-14 PROTEIN PRECURSOR (SEQ ID NO:63)
- 4) gi|16303264|dbj|BAB70499.1| (AB063483) WNT14B [Homo sapiens] (SEQ ID NO:62)
- 5) gi|7106447|ref|NP_033548.1| (NM_009522) wingless-related MMTV integration site 3A [Mus musculus] (SEQ ID NO:67)
- 6) gi|5821261|dbj|BAA83743.1| (AB024080) Wnt-3a [Gallus gallus] (SEQ ID NO:68)

5

10

15

20

25

30

35

40

45

50

55

60

65

70

```

10      20      30      40      50
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....MALLYSSTGVVCT.....CSPEYFGLTGSFPLTILF|
g1|3915306|.....MLDGSPLARWLAAAFGLTLLLAALR.....PFAAYFGLTGSFPLTILF|
g1|16303264|.....MALLRLRLGLLACTP.....RPSAAVFGLTGSFPLTILF|
g1|7106447|.....MRPPPALALCLQLALP.....AAAYSFGLTGSFPLTILF|
g1|5821261|.....MPLFLGVLTLCLR.....LKQA.....LG.....SYPLTWS|
.....MKSFCSEVIAKRLGLKQWGCWGTPMCANKKWTSSQRSLSLW|

60      70      80      90      100
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....LTLEP.....PAAACAHYKACDRLR.....LKKKRMRCRRDPCVAETLVEAVS|
g1|3915306|.....LTLEP.....PAAACAHYKACDRLR.....LKKKRMRCRRDPCVAETLVEAVS|
g1|16303264|.....LTSEMERAYKAYKVCDDLK.....LKKKRMRCRRDPCVAETLVEAVS|
g1|7106447|.....LGTAAAPAGGGAHLKCTLLR.....LSRRGKOLCRREPCVAETLVEAVS|
g1|5821261|.....LAVGPQYSSLSTQPIIGASIPGIVPKCLRFGRNYVEIMPSVAEGVK|
DVGSIAAGHQYSSLGTQPIIGSSIPGIVPKCLRFGRNYVEIMPSVAEGVK|

110     120     130     140     150
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....YSALECCQHFGRFRWNCTLEGR.....YBASLLKRGFKETAFLYAISSAGLT|
g1|3915306|.....YSALECCQHFGRFRWNCTLEGR.....YBASLLKRGFKETAFLYAISSAGLT|
g1|16303264|.....YSALECCQHFGRFRWNCTLEGR.....YBASLLKRGFKETAFLYAISSAGLT|
g1|7106447|.....YGLLECCQHFGRFRWNCTLEGR.....YBASLLKRGFKETAFLYAISSAGLT|
g1|5821261|.....YGLLECCQHFGRFRWNCTLEGR.....YBASLLKRGFKETAFLYAISSAGLT|

160     170     180     190     200
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....HALANACSAGRMERCTCDEAPDLLENREACWGGCGSDYDFCGMVSRFAD|
g1|3915306|.....HALANACSAGRMERCTCDEAPDLLENREACWGGCGSDYDFCGMVSRFAD|
g1|16303264|.....HALANACSAGRMERCTCDEAPDLLENREACWGGCGSDYDFCGMVSRFAD|
g1|7106447|.....HALANACSAGRMERCTCDEAPDLLENREACWGGCGSDYDFCGMVSRFAD|
g1|5821261|.....HALANACSAGRMERCTCDEAPDLLENREACWGGCGSDYDFCGMVSRFAD|

210     220     230     240     250
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....ARENPPDARSAMRHNNNEAGRVITKAGVETTCCKCHGVSGSCTVRTCWRQL|
g1|3915306|.....ARENPPDARSAMRHNNNEAGRVITKAGVETTCCKCHGVSGSCTVRTCWRQL|
g1|16303264|.....ARENPPDARSAMRHNNNEAGRVITKAGVETTCCKCHGVSGSCTVRTCWRQL|
g1|7106447|.....ARENPPDARSAMRHNNNEAGRVITKAGVETTCCKCHGVSGSCTVRTCWRQL|
g1|5821261|.....ARENPPDARSAMRHNNNEAGRVITKAGVETTCCKCHGVSGSCTVRTCWRQL|

260     270     280     290     300
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....APFHEVGLHLRKHYESALKVCSITNEAAGEAGAI.....SPPRGRASGAGGSDP|
g1|3915306|.....APFHEVGLHLRKHYESALKVCSITNEAAGEAGAI.....SPPRGRASGAGGSDP|
g1|16303264|.....APFHEVGLHLRKHYESALKVCSITNEAAGEAGAI.....SPPRGRASGAGGSDP|
g1|7106447|.....APFHEVGLHLRKHYESALKVCSITNEAAGEAGAI.....SPPRGRASGAGGSDP|
g1|5821261|.....APFHEVGLHLRKHYESALKVCSITNEAAGEAGAI.....SPPRGRASGAGGSDP|

310     320     330     340     350
NOV8    .....|.....|.....|.....|.....|
g1|15082261|.....PRPDELVLVLDSPSPFLAC.....RPSPTGTAQRHREK.....QSSICCGRC|
g1|3915306|.....PRPDELVLVLDSPSPFLAC.....RPSPTGTAQRHREK.....QSSICCGRC|
g1|16303264|.....PRPDELVLVLDSPSPFLAC.....RPSPTGTAQRHREK.....QSSICCGRC|
g1|7106447|.....PRPDELVLVLDSPSPFLAC.....RPSPTGTAQRHREK.....QSSICCGRC|
g1|5821261|.....PRPDELVLVLDSPSPFLAC.....RPSPTGTAQRHREK.....QSSICCGRC|

360     370     380
NOV8    .....|.....|.....|
g1|15082261|.....HNTQSRVITRECCOVHCCYVEGROCTGREBYVTCG|
g1|3915306|.....HNTQSRVITRECCOVHCCYVEGROCTGREBYVTCG|
g1|16303264|.....HNTQSRVITRECCOVHCCYVEGROCTGREBYVTCG|
g1|7106447|.....HNTQSRVITRECCOVHCCYVEGROCTGREBYVTCG|
g1|5821261|.....HNTQSRVITRECCOVHCCYVEGROCTGREBYVTCG|

```


	P25098	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	120
	Q99LL8	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	118
	P26817	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	120
5	NOV9	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQLRGDVFOKFIESDKFTRFCQWKNV	180
	A53791	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQLRGDVFOKFIESDKFTRFCQWKNV	180
	P25098	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQLRGDVFOKFIESDKFTRFCQWKNV	180
	Q99LL8	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQLRGDVFOKFIESDKFTRFCQWKNV	178
	P26817	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQLRGDVFOKFIESDKFTRFCQWKNV	180
10	NOV9	ELNIHLTMNDFS VHRIIGRGGFGEVYGC RKADTGKMYAMKCLDKKRIKMKQGETLALNER	240
	A53791	ELNIHLTMNDFS VHRIIGRGGFGEVYGC RKADTGKMYAMKCLDKKRIKMKQGETLALNER	240
	P25098	ELNIHLTMNDFS VHRIIGRGGFGEVYGC RKADTGKMYAMKCLDKKRIKMKQGETLALNER	240
	Q99LL8	ELNIHLTMNDFS VHRIIGRGGFGEVYGC RKADTGKMYAMKCLDKKRIKMKQGETLALNER	238
15	P26817	ELNIHLTMNDFS VHRIIGRGGFGEVYGC RKADTGKMYAMKCLDKKRIKMKQGETLALNER	240
	NOV9	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	A53791	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	P25098	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
20	Q99LL8	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	298
	P26817	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	NOV9	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRI SDLGLACDFS KKKPHASVGTGHYMAPE	360
	A53791	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRI SDLGLACDFS KKKPHASVGTGHYMAPE	360
25	P25098	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRI SDLGLACDFS KKKPHASVGTGHYMAPE	360
	Q99LL8	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRI SDLGLACDFS KKKPHASVGTGHYMAPE	358
	P26817	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRI SDLGLACDFS KKKPHASVGTGHYMAPE	360
	NOV9	VLQKGVA YDSSADWFS LGCMFLKLLRGHSPFRQHKT KDKHEIDRMTLTMAVELPDSFSPE	420
30	A53791	VLQKGVA YDSSADWFS LGCMFLKLLRGHSPFRQHKT KDKHEIDRMTLTMAVELPDSFSPE	420
	P25098	VLQKGVA YDSSADWFS LGCMFLKLLRGHSPFRQHKT KDKHEIDRMTLTMAVELPDSFSPE	420
	Q99LL8	VLQKGVA YDSSADWFS LGCMFLKLLRGHSPFRQHKT KDKHEIDRMTLTMAVELPDSFSPE	418
	P26817	VLQKGVA YDSSADWFS LGCMFLKLLRGHSPFRQHKT KDKHEIDRMTLTMAVELPDSFSPE	420
35	NOV9	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKES PFFRSLDWQMVFLOKYPPPLIPPRGEVNAA	480
	A53791	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKES PFFRSLDWQMVFLOKYPPPLIPPRGEVNAA	480
	P25098	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKES PFFRSLDWQMVFLOKYPPPLIPPRGEVNAA	480
	Q99LL8	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKES PFFRSLDWQMVFLOKYPPPLIPPRGEVNAA	478
	P26817	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKES PFFRSLDWQMVFLOKYPPPLIPPRGEVNAA	480
40	NOV9	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	519
	A53791	DAFDIGSFDEEDTKGIRLLDSQELRYNFPPLTISERWQOEVAETVFDTINAETDRLEARK	540
	P25098	DAFDIGSFDEEDTKGIRLLDSQELRYNFPPLTISERWQOEVAETVFDTINAETDRLEARK	540
	Q99LL8	DAFDIGSFDEEDTKGIRLLDSQELRYNFPPLTISERWQOEVAETVFDTINAETDRLEARK	538
45	P26817	DAFDIGSFDEEDTKGIRLLDSQELRYNFPPLTISERWQOEVAETVFDTINAETDRLEARK	540
	NOV9	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	579
	A53791	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
	P25098	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
50	Q99LL8	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	598
	P26817	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
	NOV9	LTMEEQSV EETQIKERKCLLLKIRGGKQF ILQCDSDPELVQWKELRDAYREAQQLVQR	639
	A53791	LTMEEQSV EETQIKERKCLLLKIRGGKQF ILQCDSDPELVQWKELRDAYREAQQLVQR	660
55	P25098	LTMEEQSV EETQIKERKCLLLKIRGGKQF ILQCDSDPELVQWKELRDAYREAQQLVQR	660
	Q99LL8	LTMEEQSV EETQIKERKCLLLKIRGGKQF ILQCDSDPELVQWKELRDAYREAQQLVQR	658
	P26817	LTMEEQSV EETQIKERKCLLLKIRGGKQF ILQCDSDPELVQWKELRDAYREAQQLVQR	660
	NOV9	VPKMKNKPRSPVVELSKVPLVQRGSANGL	668
60	A53791	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689
	P25098	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689
	Q99LL8	VPKMKNKPRSPVVELSKVPLVQRGSANGL	687
	P26817	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689

65 Tables 9E-9L list the domain descriptions from DOMAIN analysis results against NOV9. This indicates that the NOV9 sequence has properties similar to those of other proteins known to contain this domain.

Table 10C. BLAST results for NOV10

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
ptnr:SP TREMBL-ACC:Q9UL64	ALPHA-MANNOSIDASE 6A8B - Homo sapiens	1062	763/771 (99%)	767/771 (99%)	0.0
ptnr:SP TREMBL-ACC:Q9NTJ4	HYPOTHETICAL 115.8 KDA PROTEIN - Homo sapiens	1040	715/722 (99%)	718/722 (99%)	0.0
ptnr:TREMBLNEW-ACC:AAH16253	SIMILAR TO MANNOSIDASE, ALPHA, CLASS 2C, MEMBER 1	1039	635/730 (89%)	692/730 (94%)	0.0
ptnr:SWISSPROT-ACC:P21139	Alpha-mannosidase (EC 3.2.1.24)	1040	625/731 (85%)	661/731 (90%)	0.0
ptnr:SP TREMBL-ACC:Q13358	ALPHA-MANNOSIDASE - Homo sapiens	425	425/425 (100%)	425/425 (100%)	0.0

The homology between these and other sequences is shown graphically in the ClustalW analysis shown in Table 10D. In the ClustalW alignment of the NOV10 protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 10D. ClustalW Analysis of NOV10

1)	NOV10 (SEQ ID NO:26)
2)	ptnr: ALPHA-MANNOSIDASE 6A8B - Homo sapiens (SEQ ID NO:73)
3)	ptnr: HYPOTHETICAL 115.8 KDA PROTEIN - Homo sapiens (SEQ ID NO:74)
4)	ptnr: SIMILAR TO MANNOSIDASE, ALPHA, CLASS 2C, MEMBER 1 (SEQ ID NO:75)
5)	ptnr: Alpha-mannosidase (EC 3.2.1.24) (SEQ ID NO:76)
NOV10	MAAAPFLKHHRTTTFERVEKFSVPIYFTDNLRGRLFGASCSPVAVLSSFLTHRLPYCEAVQRE-----FRPAQVG70
Q9UL64	MAAAPFLKHHRTTTFERVEKFSVPIYFTDNLRGRLFGASCSPVAVLSSFLTHRLPYCEAVQRE-----FRPAQVG70
Q9NTJ4	MAAAPFLKHHRTTTFERVEKFSVPIYFTDNLRGRLFGASCSPVAVLSSFLTHRLPYCEAVQRE-----FRPAQVG70
AAH16253	MAAAPFLKHHRTTTFERVEKFSVPIYFTDNLRGRLFGASCSPVAVLSSFLTHRLPYCEAVQRE-----FRPAQVG69
P21139	MAAAPFLKHHRTTTFERVEKFSVPIYFTDNLRGRLFGASCSPVAVLSSFLTHRLPYCEAVQRE-----FRPAQVG79
NOV10	DSFGPTWNTCWFRVELTIPKAVVQGEVHLCWE-SDGEGLVWRE-GEVQGLTH-----EGEKTSYVLIDRLGERDPP140
Q9UL64	DSFGPTWNTCWFRVELTIPKAVVQGEVHLCWE-SDGEGLVWRE-GEVQGLTH-----EGEKTSYVLIDRLGERDPP140
Q9NTJ4	DSFGPTWNTCWFRVELTIPKAVVQGEVHLCWE-SDGEGLVWRE-GEVQGLTH-----EGEKTSYVLIDRLGERDPP140
AAH16253	DSFGPTWNTCWFRVELTIPKAVVQGEVHLCWE-SDGEGLVWRE-GEVQGLTH-----EGEKTSYVLIDRLGERDPP139
P21139	DSFGPTWNTCWFRVELTIPKAVVQGEVHLCWE-SDGEGLVWRE-GEVQGLTH-----EGEKTSYVLIDRLGERDPP159
NOV10	S-LTLYVEVAGNGLGAGKGSK-IAAPDPEKMF-QLSRAELAVF-----HRDVHMLLVE-LELLLGIAF-----200
Q9UL64	S-LTLYVEVAGNGLGAGKGSK-IAAPDPEKMF-QLSRAELAVF-----HRDVHMLLVE-LELLLGIAKGLGNDQNSFC211
Q9NTJ4	S-LTLYVEVAGNGLGAGKGSK-IAAPDPEKMF-QLSRAELAVF-----HRDVHMLLVE-LELLLGIAKGLGNDQNSFC211
AAH16253	S-LTLYVEVAGNGLGAGKGSK-IAAPDPEKMF-QLSRAELAVF-----HRDVHMLLVE-LELLLGIAKGLGNDQNSFC210
P21139	S-LTLYVEVAGNGLGAGKGSK-IAAPDPEKMF-QLSRAELAVF-----HRDVHMLLVE-LELLLGIAKGLGNDQNSFC229
NOV10	ALVTANOMVAVCDPAOPETIPVACALASFFGCHGGSQHTIHAGCHCHIDTAWLWPFKQTVRKCARSWVFAQLVWRNP200
Q9UL64	ALVTANOMVAVCDPAOPETIPVACALASFFGCHGGSQHTIHAGCHCHIDTAWLWPFKQTVRKCARSWVFAQLVWRNP291
Q9NTJ4	ALVTANOMVAVCDPAOPETIPVACALASFFGCHGGSQHTIHAGCHCHIDTAWLWPFKQTVRKCARSWVFAQLVWRNP291
AAH16253	ALVTANOMVAVCDPAOPETIPVACALASFFGCHGGSQHTIHAGCHCHIDTAWLWPFKQTVRKCARSWVFAQLVWRNP290
P21139	ALVTANOMVAVCDPAOPETIPVACALASFFGCHGGSQHTIHAGCHCHIDTAWLWPFKQTVRKCARSWVFAQLVWRNP229
NOV10	-----AQQLRWKSRYPGLYSRIQEPACRGQFVPVGGTWVENDNLPSGEAMVROFLOGNFFLQEFGRHCSEFWLF272
Q9UL64	EFIFACSDAQQLRWKSRYPGLYSRIQEPACRGQFVPVGGTWVENDNLPSGEAMVROFLOGNFFLQEFGRHCSEFWLF371

	Q9NTJ4	SFIFACSGAAQQLWVKSRYPCGLYSRIQEFACRCGFVPGVGGTVNDGCLPSCGAMWROFLQCGFFLOEFQNKCKSEFWLL	371
	AAH16253	DFIFACSGAAQQLWVKSRYPCGLYSRIQEFACRCGFVPGVGGTVNDGCLPSCGAMWROFLQCGFFLOEFQNKCKSEFWLL	370
	P21139	-----TLGEGDQPSQAAALY-----ANSLVNGCDP--AQQPTSTFEEALFASQNTAB-----EGDQRCG-----	281
5	NOV10	DTFGYSAQLPQIMHCCGIRRFLLTKQLSKNLVNSFPHHFTFWFEGLDGSRVLVHFPPGDSYCMGCSVEEVLKTVNNRDKGH	352
	Q9UL64	DTFGYSAQLPQIMHCCGIRRFLLTKQLSKNLVNSFPHHFTFWFEGLDGSRVLVHFPPGDSYCMGCSVEEVLKTVNNRDKGH	451
	Q9NTJ4	DTFGYSAQLPQIMHCCGIRRFLLTKQLSKNLVNSFPHHFTFWFEGLDGSRVLVHFPPGDSYCMGCSVEEVLKTVNNRDKGH	451
	AAH16253	DTFGYSAQLPQIMHCCGIRRFLLTKQLSKNLVNSFPHHFTFWFEGLDGSRVLVHFPPGDSYCMGCSVEEVLKTVNNRDKGH	450
	P21139	-----GSSQHTLQATGCHHILDTANL-----P-----FQDTRKQNS-----	314
10	NOV10	ANHSAPLFCFGDGGCGPTOTMLDRLKRLSNTDGLPRVOLSSPRLPSALESDSEQLCTVWGELFLFELHNGTYTTTHAQIKH	432
	Q9UL64	ANHSAPLFCFGDGGCGPTOTMLDRLKRLSNTDGLPRVOLSSPRLPSALESDSEQLCTVWGELFLFELHNGTYTTTHAQIKH	531
	Q9NTJ4	ANHSAPLFCFGDGGCGPTOTMLDRLKRLSNTDGLPRVOLSSPRLPSALESDSEQLCTVWGELFLFELHNGTYTTTHAQIKH	531
	AAH16253	ANHSAPLFCFGDGGCGPTOTMLDRLKRLSNTDGLPRVOLSSPRLPSALESDSEQLCTVWGELFLFELHNGTYTTTHAQIKH	530
	P21139	-----TNHSGFLFCFGDGGCGPTOTMLDRLKRLSNTDGLPRVOLSSPRLPSALESDSEQLCTVWGELFLFELHNGTYTTTHAQIKH-----	325
15	NOV10	GNRECEIRLHDVELLSLALARSAPFLYPAALQHLWRLLLLAFHDVVTGSCIONWABEAMCHYEDIRSHGNTLLSAAV	512
	Q9UL64	GNRECEIRLHDVELLSLALARSAPFLYPAALQHLWRLLLLAFHDVVTGSCIONWABEAMCHYEDIRSHGNTLLSAAV	611
	Q9NTJ4	GNRECEIRLHDVELLSLALARSAPFLYPAALQHLWRLLLLAFHDVVTGSCIONWABEAMCHYEDIRSHGNTLLSAAV	611
	AAH16253	GNRECEIRLHDVELLSLALARSAPFLYPAALQHLWRLLLLAFHDVVTGSCIONWABEAMCHYEDIRSHGNTLLSAAV	610
	P21139	-----RN-----TPTFTFCGCGCTAB-----ETQLEWVK-----NGLP-----EG-----LYAQLEGFAN-----	365
20	NOV10	AALCAGEPGPEGILLVNTLPWKRIEYMALPKPGGAHSLAULTVPSMGYAPVPPTSLQPLLPOOPPVFVVOETCGSVTLN	592
	Q9UL64	AALCAGEPGPEGILLVNTLPWKRIEYMALPKPGGAHSLAULTVPSMGYAPVPPTSLQPLLPOOPPVFVVOETCGSVTLN	691
	Q9NTJ4	AALCAGEPGPEGILLVNTLPWKRIEYMALPKPGGAHSLAULTVPSMGYAPVPPTSLQPLLPOOPPVFVVOETCGSVTLN	691
	AAH16253	AALCAGEPGPEGILLVNTLPWKRIEYMALPKPGGAHSLAULTVPSMGYAPVPPTSLQPLLPOOPPVFVVOETCGSVTLN	690
	P21139	-----AALCAGEPGPEGILLVNTLPWKRIEYMALPKPGGAHSLAULTVPSMGYAPVPPTSLQPLLPOOPPVFVVOETCGSVTLN-----	374
25	NOV10	GIIIRVKLDPTGRLTSLVLVASGREAIAGCAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGEGGLRGS	672
	Q9UL64	GIIIRVKLDPTGRLTSLVLVASGREAIAGCAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGEGGLRGS	771
	Q9NTJ4	GIIIRVKLDPTGRLTSLVLVASGREAIAGCAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGEGGLRGS	771
	AAH16253	GIIIRVKLDPTGRLTSLVLVASGREAIAGCAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGEGGLRGS	770
	P21139	-----GIIIRVKLDPTGRLTSLVLVASGREAIAGCAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGEGGLRGS-----	402
30	NOV10	WFLLOISNRSRLSQEVLDVGCVPVRFHTEVHHWHAHFKLVKEFPAPRVSSQATYBICFCHLQRPPTHYNTSDHAKRFEVK	752
	Q9UL64	WFLLOISNRSRLSQEVLDVGCVPVRFHTEVHHWHAHFKLVKEFPAPRVSSQATYBICFCHLQRPPTHYNTSDHAKRFEVK	851
	Q9NTJ4	WFLLOISNRSRLSQEVLDVGCVPVRFHTEVHHWHAHFKLVKEFPAPRVSSQATYBICFCHLQRPPTHYNTSDHAKRFEVK	851
	AAH16253	WFLLOISNRSRLSQEVLDVGCVPVRFHTEVHHWHAHFKLVKEFPAPRVSSQATYBICFCHLQRPPTHYNTSDHAKRFEVK	850
	P21139	-----WFLLOISNRSRLSQEVLDVGCVPVRFHTEVHHWHAHFKLVKEFPAPRVSSQATYBICFCHLQRPPTHYNTSDHAKRFEVK-----	408
35	NOV10	AHRWMDLSEHGFGGLALNDCKYKASVGRSILSLSLRAPKAPDATADTGRHEFTYALMPHKGSFQDAGVIOAAYSUNFPI	931
	Q9UL64	AHRWMDLSEHGFGGLALNDCKYKASVGRSILSLSLRAPKAPDATADTGRHEFTYALMPHKGSFQDAGVIOAAYSUNFPI	931
	Q9NTJ4	AHRWMDLSEHGFGGLALNDCKYKASVGRSILSLSLRAPKAPDATADTGRHEFTYALMPHKGSFQDAGVIOAAYSUNFPI	931
	AAH16253	AHRWMDLSEHGFGGLALNDCKYKASVGRSILSLSLRAPKAPDATADTGRHEFTYALMPHKGSFQDAGVIOAAYSUNFPI	930
	P21139	-----AHRWMDLSEHGFGGLALNDCKYKASVGRSILSLSLRAPKAPDATADTGRHEFTYALMPHKGSFQDAGVIOAAYSUNFPI-----	408
40	NOV10	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	912
	Q9UL64	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1011
	Q9NTJ4	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1011
	AAH16253	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1010
	P21139	-----LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH-----	408
45	NOV10	LTSGOPPBAHLFSLPSAVPPLARSASAILSPYGNWFCRRRLNGLLLSASFA- 963	
	Q9UL64	LTSGOPPBAHLFSLPSAVPPLARSASAILSPYGNWFCRRRLNGLLLSASFA- 1062	
	Q9NTJ4	LTSGOPPBAHLFSLPSAVPPLARSASAILSPYGNWFCRRRLNGLLLSASFA- 1040	
	AAH16253	LTSGOPPBAHLFSLPSAVPPLARSASAILSPYGNWFCRRRLNGLLLSASFA- 1039	
	P21139	-----LTSGOPPBAHLFSLPSAVPPLARSASAILSPYGNWFCRRRLNGLLLSASFA-----	408
50	NOV10	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	912
	Q9UL64	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1011
	Q9NTJ4	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1011
	AAH16253	LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH	1010
	P21139	-----LALPAPSPAPATSWSAFVSPPAVVLETVKQAESSPCRRSLVLRLEYAAGSHVDCVHLHLSLPVQEAILLCDLLERPDAGH-----	408
55			

60 Table 10E lists the domain description from DOMAIN analysis results against NOV10. This indicates that the NOV10 sequence has properties similar to those of other proteins known to contain this domain.

Model	Description	Score	E-value
Glyco_hydro_38 (InterPro)	Glycosyl hydrolases family 38	140.5	1e-39
(SEQ ID NO:111)			
Glyco_hydro_38: domain 1 of 2, from 230 to 332: score 89.2, E = 5.4e-25			
AC058790_d	230	VGGTGWEMDGNLPSGEAMVRQFLQGQNFLLQEFQ--KMCSEFWLPDT	274
AC058790_d	275	FGHSatmPyLlraqaGfdgflIqRihYadKksfaetkqleFvWrqswslt FGYSAQLPQIM-HGCGIRRFLTQKLSWNVLNSFPHHT---PFWF---GLD	317

Q9ZSN4 AIAQADQNYDYASNSVILHLDAGDEVFIKLDGGKRAHGGNSNKYSTFSGFIIYSD--- 258
P02746 TFCDYAYNTFQVTIGGMVLEKLEGGENVFLCATDKNSLLGMEGANSIFSGFLLFPDMEA 251
AAH08983 TFCDYAYNTFQVTIGGMVLEKLEGGENVFLCATDKNSLLGMEGANSIFSGFLLFPDMEA 253

5

Tables 11E-11F list the domain descriptions from DOMAIN analysis results against NOV11. This indicates that the NOV11 sequence has properties similar to those of other proteins known to contain this domain.

10

Table 11E. Domain Analysis of NOV11

gnl|Smart|smart00110, Clq, Complement component Clq domain.; Globular domain found in many collagens and eponymously in complement Clq. When part of full length proteins these domains form a 'bouquet' due to the multimerization of heterotrimers. The Clq fold is similar to that of tumour necrosis factor. (SEQ ID NO:104)
CD-Length = 132 residues, 99.2% aligned
Score = 113 bits (283), Expect = 1e-26

Query: 108 PRIAFYAGL--RRPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGG 165
|| || || + + ||| + | | + + ||||| + ||| + + || +
15 Sbjct: 2 PRSAPSVIRSTNRPPPGQPVRFDKVLNQGHYDPSTGKFTCPVPGVYFFSYHIESK-- 59
Query: 166 DGTSMWADLMKNGQVRASAIQADQNYDYASNSVILHLDVGDEVFIKLDGGKVHG--GNT 224
| ++ ||||| + | || + | ||| + + + || |
20 Sbjct: 60 -GRNVKVSMLKNGIQVMRECDEYQKGLYQVASGGALLQLRQGDQVWLELDDKKNGLYAGE 118
Query: 225 NKYSTFSGFIIYPD 238
||||| + + + ||
Sbjct: 119 EVDSTFSGFLLFPD 132

25

Table 11F. Domain Analysis of NOV11

gnl|Pfam|pfam00386, Clq, Clq domain. Clq is a subunit of the C1 enzyme complex that activates the serum complement system. (SEQ ID NO:112)
CD-Length = 125 residues, 100.0% aligned
Score = 102 bits (253), Expect = 3e-23

Query: 111 AFYAGLR-RPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTS 169
|| | || + + ||| + | | + + ||||| + ||| + + || + || +
30 Sbjct: 1 AFTAIRSTRPPAPGQPVIFDEVLYNQGHYDPATGKFTCPVPGLYYFNHVSCK--GTN 57
Query: 170 MWADLMKNGQVRASAIQADQNYDYASNSVILHLDVGDEVFIKLDGGKVHG--GNTNKY 227
+ ||| + | + | | || + | ||| + + + || + + | | +
Sbjct: 58 VCVSLMRNGVPVMSFCDEYAKGTYQVASGGAQLRQGDQVWLELDDKQTNGLLGEGVH 117
Query: 228 STFSGFII 235
| |||| + +
35 Sbjct: 118 SVFSGFLL 125

The first component of complement system is a calcium-dependent complex of the 3 subcomponents Clq, Clr, and C1s. Subcomponent Clq binds to immunoglobulin complexes

40

ptnr: SPTREMBL- ACC: P70208	PLEXIN 3 - Mus musculus	1872	1245/1874 (66%)	1478/1874 (78%)	0.0
--------------------------------	----------------------------	------	--------------------	--------------------	-----

The homology between these and other sequences is shown graphically in the ClustalW analysis shown in Table 12D. In the ClustalW alignment of the NOV12 protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 12D. ClustalW Analysis of NOV12

10	1) NOV12 (SEQ ID NO:32)
	2) ptnr: PLEXIN 1 - Mus musculus (SEQ ID NO:81)
	3) ptnr: NOV/PLEXIN-A1 PROTEIN - Homo sapiens (SEQ ID NO:82)
	4) ptnr: PLEXIN PRECURSOR - Xenopus laevis (SEQ ID NO:83)
15	5) ptnr: Plexin A3 precursor (Plexin 4) (SEQ ID NO:84)
6)	
	NOV12 MMLTPAQPEHRGPRPQAMPDPPPSLQVLLLLLLLLLPQVABASLFRACCGSOPPFRTESASDWGLTHLVVHEOTGR 80
20	P70206 -----PLPLPLSSRTLLLLLLLRVYHITISSPPAGLGPQAFRTFVASDWGLTHLVVHEOTGE 60
	Q9UIW2 -----LAAEASLFRACCGSOPPFRTESASDWGLTHLVVHEOTGE 40
	Q91823 -----LLHAERPLPPIHMTFVLLGSAHPT-----SDGSPKDFRTFASDWGLTHLVVHNTIGE 55
	P51805 -----LPSVCLLLPLPLV-----LALGNRPFSAEVVEPTLTHLVHVRVIGE 44
25	NOV12 VYVGAVNRIYKLSQNLTLRAHVTCGPVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 160
	P70206 VYVGAVNRIYKLSQNLTLRAHVTCGPVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 140
	Q9UIW2 VYVGAVNRIYKLSQNLTLRAHVTCGPVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 120
	Q91823 VYVGAVNRIYKLSQNLTLRAHVTCGPVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYSDNRIACGSASQGI COFLR 135
	P51805 VYVGAVNRIYKLSQNLTLRAHVTCGPVEDNRCYPPPSVAVLAHRLAPVDVANKLLLDYAAARLLACGSASQGI COFLR 124
30	NOV12 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMFQFVYQ 240
	P70206 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMFQFVYQ 220
	Q9UIW2 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMFQFVYQ 200
	Q91823 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMFQFVYQ 160
35	P51805 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMFQFVYQ 202
40	NOV12 EFVSSOLKI PSOTLSKFPADFIVVYVSRSEQFVYVLTQLDQTLTSPDAAGEHFTSKIVRLCVDOPKFYSVVEFPICG 320
	P70206 EFVSSOLKI PSOTLSKFPADFIVVYVSRSEQFVYVLTQLDQTLTSPDAAGEHFTSKIVRLCVDOPKFYSVVEFPICG 300
	Q9UIW2 EFVSSOLKI PSOTLSKFPADFIVVYVSRSEQFVYVLTQLDQTLTSPDAAGEHFTSKIVRLCVDOPKFYSVVEFPICG 280
	Q91823 EFVSSOLKI PSOTLSKFPADFIVVYVSRSEQFVYVLTQLDQTLTSPDAAGEHFTSKIVRLCVDOPKFYSVVEFPICG 160
	P51805 EFVSSOLKI PSOTLSKFPADFIVVYVSRSEQFVYVLTQLDQTLTSPDAAGEHFTSKIVRLCVDOPKFYSVVEFPICG 282
45	NOV12 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOGGPNRVKPPKESALCLFLTRAIKEKIKERIOSCYRGECH 400
	P70206 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOGGPNRVKPPKESALCLFLTRAIKEKIKERIOSCYRGECH 380
	Q9UIW2 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOGGPNRVKPPKESALCLFLTRAIKEKIKERIOSCYRGECH 360
	Q91823 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOGGPNRVKPPKESALCLFLTRAIKEKIKERIOSCYRGECH 160
	P51805 SRRGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOGGPNRVKPPKESALCLFLTRAIKEKIKERIOSCYRGECH 362
50	NOV12 LSLPWLINKELGCCINSPLQIDDDFCQDQNPOLGQTVTIECTPLFVDRDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 480
	P70206 LSLPWLINKELGCCINSPLQIDDDFCQDQNPOLGQTVTIECTPLFVDRDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 460
	Q9UIW2 LSLPWLINKELGCCINSPLQIDDDFCQDQNPOLGQTVTIECTPLFVDRDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 440
	Q91823 LSLPWLINKELGCCINSPLQIDDDFCQDQNPOLGQTVTIECTPLFVDRDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 160
	P51805 LSLPWLINKELGCCINSPLQIDDDFCQDQNPOLGQTVTIECTPLFVDRDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 442
55	NOV12 DLSNFGRRPALAYESVVAOEGSPILRDLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 560
	P70206 DLSNFGRRPALAYESVVAOEGSPILRDLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 540
	Q9UIW2 DLSNFGRRPALAYESVVAOEGSPILRDLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 520
	Q91823 DLSNFGRRPALAYESVVAOEGSPILRDLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 160
	P51805 EG-----PDSNHLVTVPVVVGSPILRDLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 518
60	NOV12 RQDACEARAEPORFASDLIQCQVLTQVQPNVSVVTMSQVFLVLQANVVDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 639
	P70206 RQDACEARAEPORFASDLIQCQVLTQVQPNVSVVTMSQVFLVLQANVVDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 618
	Q9UIW2 RQDACEARAEPORFASDLIQCQVLTQVQPNVSVVTMSQVFLVLQANVVDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 598
	Q91823 RQDACEARAEPORFASDLIQCQVLTQVQPNVSVVTMSQVFLVLQANVVDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 160
65	P51805 RQDACEARAEPORFASDLIQCQVLTQVQPNVSVVTMSQVFLVLQANVVDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 597
	NOV12 SAREVAPITRCCGEGSDQRVVKLYLKSKEKTKGFASVDVFVYVCSVHOSGLSCVNGSFPCHWCKYRHVCTHVVADCAFLR 719
	P70206 SAREVAPITRCCGEGSDQRVVKLYLKSKEKTKGFASVDVFVYVCSVHOSGLSCVNGSFPCHWCKYRHVCTHVVADCAFLR 695

	Q9UIW2	SAREVAPITRGC--GDORVVKLYNKSKEGCKKFAVDVFVNCVHCS	CLSCVNGSPFCHCKYRHVC	DNVADCAFL	675
	Q91823	-----	-----	-----	160
	P51805	SLGGRALTRGH--SATGTVLCLSLKSTGVRFADDFVFNCSVLAS	SLGCVOSNPPCHCKYRHVC	SRPHCS	674
5	NOV12	GRWVSEDCPOLPSTQIYVPVGVVKPITLAAPNLPOPOSGQRCGYECLFHIPGSPARVTALRFN	SSSLOCCN	SSSYEGN	799
	P70206	GRWVSEDCPOLPSTQIYVPVGVVKPITLAAPNLPOPOSGQRCGYECLFHIPGSPARVTALRFN	SSSLOCCN	SSSYEGN	775
	Q9UIW2	GRWVSEDCPOLPSTQIYVPVGVVKPITLAAPNLPOPOSGQRCGYECLFHIPGSPARVTALRFN	SSSLOCCN	SSSYEGN	755
	Q91823	-----	-----	-----	160
	P51805	GRWVSEDCPOLPSTQIYVPVGVVKPITLAAPNLPOPOSGQRCGYECLFHIPGSPARVTALRFN	SSSLOCCN	SSSYEGN	754
10	NOV12	DVSDLPVNLVVVKNFVINDPQNIQAHLYKCPALPSCGLCLKADPRFECGWVAERKCSLRHHIC	ADTPASMKHARIC		879
	P70206	DVSDLPVNLVVVKNFVINDPQNIQAHLYKCPALPSCGLCLKADPRFECGWVAERKCSLRHHIC	ADTPASMKHARIC		855
	Q9UIW2	DVSDLPVNLVVVKNFVINDPQNIQAHLYKCPALPSCGLCLKADPRFECGWVAERKCSLRHHIC	ADTPASMKHARIC		835
	Q91823	-----	-----	-----	160
	P51805	HHGCTEDFVVVNDGDFTEKPSFPAALYKCAQAPSCGLCLKADPRFECGWVAERKCSLRHHIC	ADTPASMKHARIC		833
15	NOV12	SSRCTDPKILKLSPEGTGPGCGTALTTTCENLGLRPFEDVRLGVVGVKVLCSPVSEFYISAEQIVCEIGDASSVRAHDALV			959
	P70206	SSRCTDPKILKLSPEGTGPGCGTALTTTCENLGLRPFEDVRLGVVGVKVLCSPVSEFYISAEQIVCEIGDASSVRAHDALV			935
	Q9UIW2	SSRCTDPKILKLSPEGTGPGCGTALTTTCENLGLRPFEDVRLGVVGVKVLCSPVSEFYISAEQIVCEIGDASSVRAHDALV			915
	Q91823	-----	-----	-----	160
	P51805	CHQSHDRIITTHGLVGEKCGCTFTIVGEMGLLSRSTVGRAG--DPCNSHPAEYISAEQIVCEIGDASSVRAHDALV			911
20	NOV12	EVCVADCSFHYRALSPPKRTFTVTPTFYRVSPSRGLSGGTWIGIEGSHINAGSOVAVSVGGRCPSFSKSRP	NSREIRCLT		1039
	P70206	EVCVADCSFHYRALSPPKRTFTVTPTFYRVSPSRGLSGGTWIGIEGSHINAGSOVAVSVGGRCPSFSKSRP	NSREIRCLT		1013
	Q9UIW2	EVCVADCSFHYRALSPPKRTFTVTPTFYRVSPSRGLSGGTWIGIEGSHINAGSOVAVSVGGRCPSFSKSRP	NSREIRCLT		993
	Q91823	-----	-----	-----	160
	P51805	SLGVDCSADFSTGCEGVSEVTPTEQVSPSRGLSGGTWIGIEGSHINAGSOVAVSVGGRCPSFSKSRP	NSREIRCLT		989
25	NOV12	PPCGH--PGSAPITILINRAOITNPEVKYNYTEDPTILRIDPEWSINSGTLLTVTCTNLATVREPRIRAKYGGIEREN			1117
	P70206	PPCGH--PGSAPITILINRAOITNPEVKYNYTEDPTILRIDPEWSINSGTLLTVTCTNLATVREPRIRAKYGGIEREN			1092
	Q9UIW2	PPCGH--PGSAPITILINRAOITNPEVKYNYTEDPTILRIDPEWSINSGTLLTVTCTNLATVREPRIRAKYGGIEREN			1072
	Q91823	-----	-----	-----	160
	P51805	PLSTLGSAPITILINRAOITNPEVKYNYTEDPTILRIDPEWSINSGTLLTVTCTNLATVREPRIRAKYGGIEREN			1069
30	NOV12	LYNDITMVCRAFSVANDVSRPPELGERPDELGFVMDVRSLLVLNHSFLYYPCPVLEPLSPITGLLEKPKSSPLILKCH			1197
	P70206	LYNDITMVCRAFSVANDVSRPPELGERPDELGFVMDVRSLLVLNHSFLYYPCPVLEPLSPITGLLEKPKSSPLILKCH			1172
	Q9UIW2	LYNDITMVCRAFSVANDVSRPPELGERPDELGFVMDVRSLLVLNHSFLYYPCPVLEPLSPITGLLEKPKSSPLILKCH			1152
	Q91823	-----	-----	-----	160
	P51805	QIINDITMVCRAFSVANDVSRPPELGERPDELGFVMDVRSLLVLNHSFLYYPCPVLEPLSPITGLLEKPKSSPLILKCH			1149
35	NOV12	NLLPPAPONSRYITVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGFESPGTLQVYSDSLTLPAIVGIGGGGG			1277
	P70206	NLLPPAPONSRYITVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGFESPGTLQVYSDSLTLPAIVGIGGGGG			1252
	Q9UIW2	NLLPPAPONSRYITVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGFESPGTLQVYSDSLTLPAIVGIGGGGG			1232
	Q91823	-----	-----	-----	160
	P51805	NLLPPAPONSRYITVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGFESPGTLQVYSDSLTLPAIVGIGGGGG			1229
40	NOV12	LLLLVIVAVLIAYKKRSDADRTLKRLQLOQNDLESRVALECKEAFASLOTDIHETLNOLDCAGIPFLOYRTYAMRVLPF			1357
	P70206	LLLLVIVAVLIAYKKRSDADRTLKRLQLOQNDLESRVALECKEAFASLOTDIHETLNOLDCAGIPFLOYRTYAMRVLPF			1332
	Q9UIW2	LLLLVIVAVLIAYKKRSDADRTLKRLQLOQNDLESRVALECKEAFASLOTDIHETLNOLDCAGIPFLOYRTYAMRVLPF			1312
	Q91823	-----	-----	-----	160
	P51805	LLLLVIVAVLIAYKKRSDADRTLKRLQLOQNDLESRVALECKEAFASLOTDIHETLNOLDCAGIPFLOYRTYAMRVLPF			1309
45	NOV12	GLEDHPVLKEMEVOANVEKSLTLFGQLTKKHFLTLFIRTLAQRFSFMRDRGNVASLIMTALOGEMEYATGVLLKOLLS			1437
	P70206	GLEDHPVLKEMEVOANVEKSLTLFGQLTKKHFLTLFIRTLAQRFSFMRDRGNVASLIMTALOGEMEYATGVLLKOLLS			1412
	Q9UIW2	GLEDHPVLKEMEVOANVEKSLTLFGQLTKKHFLTLFIRTLAQRFSFMRDRGNVASLIMTALOGEMEYATGVLLKOLLS			1392
	Q91823	-----	-----	-----	160
	P51805	GLEDHPVLKEMEVOANVEKSLTLFGQLTKKHFLTLFIRTLAQRFSFMRDRGNVASLIMTALOGEMEYATGVLLKOLLS			1389
50	NOV12	LIERNLESNNHPKLLLRPFESVAEKMLTMTFTFLYKFLNECAGEPLFMYCAIKCOMKGPIDAITGEARYSLSEDKI			1517
	P70206	LIERNLESNNHPKLLLRPFESVAEKMLTMTFTFLYKFLNECAGEPLFMYCAIKCOMKGPIDAITGEARYSLSEDKI			1491
	Q9UIW2	LIERNLESNNHPKLLLRPFESVAEKMLTMTFTFLYKFLNECAGEPLFMYCAIKCOMKGPIDAITGEARYSLSEDKI			1471
	Q91823	-----	-----	-----	160
	P51805	LIERNLESNNHPKLLLRPFESVAEKMLTMTFTFLYKFLNECAGEPLFMYCAIKCOMKGPIDAITGEARYSLSEDKI			1468
55	NOV12	IRQOIDYKTLTLNCAVPEENAEVVPVAGLDCDTVTOAKEKLLDAVYKGVYSORPKADMDLEWROGMARI			1597
	P70206	IRQOIDYKTLTLNCAVPEENAEVVPVAGLDCDTVTOAKEKLLDAVYKGVYSORPKADMDLEWROGMARI			1571
	Q9UIW2	IRQOIDYKTLTLNCAVPEENAEVVPVAGLDCDTVTOAKEKLLDAVYKGVYSORPKADMDLEWROGMARI			1551
	Q91823	-----	-----	-----	160
	P51805	IRQOIDYKTLTLNCAVPEENAEVVPVAGLDCDTVTOAKEKLLDAVYKGVYSORPKADMDLEWROGMARI			1548
60	NOV12	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYESMLRTASSPDSLSRSTPHITPDLESCTKI			1677
	P70206	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYESMLRTASSPDSLSRSTPHITPDLESCTKI			1651
	Q9UIW2	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYESMLRTASSPDSLSRSTPHITPDLESCTKI			1631
	Q91823	-----	-----	-----	160
	P51805	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYESMLRTASSPDSLSRSTPHITPDLESCTKI			1628
65	NOV12	MHLVNMHDLDOREGDRGSKMVSSEIYLLRLLATKQ--FLQKFVDDLFETIFSTAHRSALPLAIKYMDFLDEQADKHQIH			1757
	P70206	MHLVNMHDLDOREGDRGSKMVSSEIYLLRLLATKQ--FLQKFVDDLFETIFSTAHRSALPLAIKYMDFLDEQADKHQIH			1730
	Q9UIW2	MHLVNMHDLDOREGDRGSKMVSSEIYLLRLLATKQ--FLQKFVDDLFETIFSTAHRSALPLAIKYMDFLDEQADKHQIH			1710
	Q91823	-----	-----	-----	160
	P51805	MHLVNMHDLDOREGDRGSKMVSSEIYLLRLLATKQ--FLQKFVDDLFETIFSTAHRSALPLAIKYMDFLDEQADKHQIH			1707
70	NOV12	PDVVRHTWKSNG--LPLRFVWVITNPOFVFDIHNISITDACLSVVAOTFMDSCSTSEHILGKDSPSNKLKAKDIYVKS			1837
	P70206	PDVVRHTWKSNG--LPLRFVWVITNPOFVFDIHNISITDACLSVVAOTFMDSCSTSEHILGKDSPSNKLKAKDIYVKS			1809
	Q9UIW2	PDVVRHTWKSNG--LPLRFVWVITNPOFVFDIHNISITDACLSVVAOTFMDSCSTSEHILGKDSPSNKLKAKDIYVKS			1754
	Q91823	-----	-----	-----	160
	P51805	PDVVRHTWKSNG--LPLRFVWVITNPOFVFDIHNISITDACLSVVAOTFMDSCSTSEHILGKDSPSNKLKAKDIYVKS			1786
75	NOV12	MVERRYADIANKPISIDODHSAYLAROSRLHSCNSGALHETISLHARWD--MILVALKCPONRORLSKLENY			1917
	P70206	MVERRYADIANKPISIDODHSAYLAROSRLHSCNSGALHETISLHARWD--MILVALKCPONRORLSKLENY			1886

Q9UIW2 1754
 Q91823 160
 P51805 1863
 NOV12 1925
 P70206 1894
 Q9UIW2 1754
 Q91823 160
 P51805 1871

Tables 12E-12N list the domain descriptions from DOMAIN analysis results against NOV12. This indicates that the NOV12 sequence has properties similar to those of other proteins known to contain this domain.

Table 12E. Domain Analysis of NOV12

gnl|Smart|smart00630, Sema, semaphorin domain (SEQ ID NO:113)
 CD-Length = 430 residues, 100.0% aligned
 Score = 242 bits (618), Expect = 1e-64

Query: 69 LTHLVVHEQTGEVYVGAVNRIYKLSGNLTLLRAHVTGPVEDNEKCYPPSPVQSCPHGLGS 128
 Sbjct: 1 LQNLLDDEDNGLTYVGARNRLYVLSLNLISEAEVKTGPVLSSPDCEEC--VSKGKDP-- 56
 Query: 129 TDNVNK-LLLLDYAANRLACGS-ASQGICQFLRLDDLFLKLGEPHHRKEHYLSSVQEAGS 186
 Sbjct: 57 TDCVNFIRLLLDYNADHLLVCGTNAFQPVCRILNLGNLDRLE-EVGRESGRGRCPDPQHN 115
 Query: 187 MAGVLIAGPPGQQAQLFVGTPID--GKSEYFPTLSSRLMANEEDADMFGFVYQDEFVS 244
 Sbjct: 116 STAVLVGD-----ELYVGTVADFGSDPAIYRSLSVRRLLKGTSG-----PSLRTVL 161
 Query: 245 SQLKIPSDTLKFPFAFDIYYVYSFRSEQFVYYLTLQLDTQLTSPDAAGEHFFTSKIVRLC 304
 Sbjct: 162 YDSRWLN-----EPNFVYAFESGDFVYF----FFRETAVEDENCGKAVVSRVARVC 208
 Query: 305 VDD-----PKFYSYVEFFIGC---EQAGVEYRLVQDAYLSRPGRALAHQLGLAEDED 353
 Sbjct: 209 KNDVGGPRSLSKKWTSLFKARLECSVPGEFFPYFNLQAAFLLPAG-----SESDD 259
 Query: 354 VLFTVFAQGQKNRVKPPKESALCLFTLRAIKEKIKERIQSCYRGEKLSL----PWLLNK 409
 Sbjct: 260 VLYGVFSTS----SNPIGSAVCAFSLSDINAVFNEPFKECETGNSQWLPYPRGLVPFPR 315
 Query: 410 ELGCINSPLQI----DDDFC-GQDFNQPLGGTVTIEGTPLFV--DKDDGLTAVAA----Y 458
 Sbjct: 316 PGTCNPNTPLSSKDLPPDVLNFIKTHPLMDEVVQPLTGRPLFVKTDNLYLLTSIAVDRVRT 375
 Query: 459 DYRGRTVVFAGTRSGRIKILVDLSNPGGRPALAYESVVAQEGSPILRDLVLSPNH 514
 Sbjct: 376 DGGNYTVLFLGTS DGRILKVLSRSSSSSES VLEEISVFDPGSPV-SDLVLSPPK 430

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.